

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 17:12:26 ; Search time 5147 Seconds  
(without alignments)  
10252.133 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcggcagcgccgca.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sta.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	88.6	1089	6 AX100704	AX100704 Sequence
2	648	59.5	1083	6 TSP271917	AJ271917 Triticum
3	168	15.4	168	6 AX100712	AX100712 Sequence
4	168	15.4	168	6 AX100714	AX100714 Sequence
5	153	14.0	153	6 AX100710	AX100710 Sequence
6	150	13.8	150	6 AX100716	AX100716 Sequence
7	114	10.5	156	6 AX100708	AX100708 Sequence
8	60	5.5	1132	8 SLU39059	U39059 Lycopersico
9	59	5.4	1630	9 BC028028	BC028028 Homo sapi
10	57	5.2	246093	2 CR628367	CR628367 Mus muscu
11	56	5.1	444	9 BC002750	BC002750 Homo sapi
12	56	5.1	1980	9 BC015791	BC015791 Homo sapi
13	56	5.1	1985	3 AX118692	AX118692 Drosophil
14	56	5.1	2084	9 AB072761	AB072761 Macaca fa
15	56	5.1	3430	9 BC065220	BC065220 Homo sapi
16	56	5.1	3441	10 BC038365	BC038365 Mus muscu
17	56	5.1	3606	9 BC030162	BC030162 Homo sapi
18	56	5.1	4430	5 BC059794	BC059794 Danio rer
19	56	5.1	4669	5 BC048022	BC048022 Xenopus 1

56 5.1 114458 10 BX293986 BX293986 Mouse DNA  
56 5.1 161155 2 AC128086 AC128086 Rattus no  
56 5.1 179678 2 AC123687 AC123687 Mus muscu  
56 5.1 208225 2 AC110537 AC110537 Mus muscu  
56 5.1 237538 2 AC108561 AC108561 Rattus no  
56 5.1 244978 2 AC128734 AC128734 Rattus no  
56 5.1 376 6 CQ526562 CQ526562 Sequence  
55 5.1 491 6 AR487921 AR487921 Sequence  
55 5.1 764 5 AF354917 AF354917 Trimeresu  
55 5.1 798 9 AK026442 AK026442 Homo sapi  
55 5.1 833 9 AK024546 AK024546 Homo sapi  
55 5.1 840 5 BC083489 BC083489 Danio rer  
55 5.1 850 9 BC036763 BC036763 Homo sapi  
55 5.1 888 9 BC046163 BC046163 Homo sapi  
55 5.1 1489 10 BC028982 BC028982 Mus muscu  
55 5.1 1655 10 BC061850 BC061850 Rattus no  
55 5.1 1723 5 BC068376 BC068376 Danio rer  
55 5.1 1789 9 AK074236 AK074236 Homo sapi  
55 5.1 2021 9 BC008075 BC008075 Homo sapi  
55 5.1 2528 3 AY051822 AY051822 Drosophil  
55 5.1 2673 10 BC021308 BC021308 Mus muscu  
55 5.1 2727 5 BC067323 BC067323 Xenopus t  
55 5.1 2859 10 BC033303 BC033303 Mus muscu  
55 5.1 2973 9 BC042518 BC042518 Homo sapi  
55 5.1 3422 9 HSM805920 HSM805920 Homo sapi  
55 5.1 3819 6 CQ768015 CQ768015 Sequence

#### ALIGNMENTS

RESULT 1  
AX100704  
LOCUS AX100704 1089 bp DNA linear PAT 10-APR-2001  
DEFINITION Sequence 1 from Patent WO0121644.  
ACCESSION AX100704  
VERSION AX100704.1 GI:13619652  
SOURCE Triticum monococcum  
ORGANISM Triticum monococcum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
REFERENCE 1  
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.  
TITLE Wheat gp proteins and uses thereof  
JOURNAL Patent: WO 0121644-A 1 29-MAR-2001;  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)  
FEATURES  
Location/Qualifiers  
source 1..1089  
/organism="Triticum monococcum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4568"  
20..805  
/codon\_start=1  
/protein\_id="CAC36466.1"  
/db\_xref="GI:13619653"  
/translation="MAPPPGAAATAALDLTGVIHLEASSVPPIPERGGNAVQRKG  
AYDPDKRKKKAAAPRITGWLREYSKIVCEKVEAKRTTYNEADEIYSELKMAH  
IQGQDEKNIRRVDAFNVLIAKEKEIRWMGLSNRYEKIKKLEEVKELV  
NKIRNKALQIEKQFDLQNIKLNOTLESSAENVNGIRLPFLVLTSTRKARVEIE  
ISDDSKFAHFENGAFFTLHDDLISLEGVRRNSIGRATLH"

#### CDS

Query Match 88.6%; Score 965; DB 6; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 GCGGGTAATGCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184  
DB 125 GCGGGTAATGCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184

#### ORIGIN

Query Match 88.6%; Score 965; DB 6; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 GCGGGTAATGCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184  
DB 125 GCGGGTAATGCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184

QY 185 AAGCTGCGGACCGAGGATCACCGTTCGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244  
Db |||||||  
QY 245 GAGAAAGTTGAAGCAGGAGGAGCAACATACAAATGAGGTTGCAGAGAAATTTATTCA 304  
Db |||||||  
QY 245 GAGAAAGTTGAAGCAGGAGGAGCAACATACAAATGAGGTTGCAGAGAAATTTATTCA 304  
Db |||||||  
QY 305 GAGCTGAAGTCCATGCGACATATTTGTTCAAGGTTTGTGAGAGAAATTTAGCGCGAGA 364  
Db |||||||  
QY 305 GAGCTGAAGTCCATGCGACATATTTGTTCAAGGTTTGTGAGAGAAATTTAGCGCGAGA 364  
Db |||||||  
QY 365 GTGTATGATGCTTTCACGTTCTCATTGACATTCGTTTATTCGAAAGAAAGAGAG 424  
Db |||||||  
QY 365 GTGTATGATGCTTTCACGTTCTCATTGACATTCGTTTATTCGAAAGAAAGAGAG 424  
Db |||||||  
QY 425 ATACGCTGGATGGGCTTTCAAAATACAGATATGAAAAATAAAGAGCTTCAGGAAGTT 484  
Db |||||||  
QY 425 ATACGCTGGATGGGCTTTCAAAATACAGATATGAAAAATAAAGAGCTTCAGGAAGTT 484  
Db |||||||  
QY 485 CGTAAAGAACTCGTCAACAGATTAGAGCAAGAGGCACTCTCCAGGAAATCGAAAAA 544  
Db |||||||  
QY 485 CGTAAAGAACTCGTCAACAGATTAGAGCAAGAGGCACTCTCCAGGAAATCGAAAAA 544  
Db |||||||  
QY 545 CAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAACCTGGAAGCTCAGCAGAG 604  
Db |||||||  
QY 545 CAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAACCTGGAAGCTCAGCAGAG 604  
Db |||||||  
QY 605 AATGTTAATGGCATCGGCTTCCATTCGTTTCAAGACATCTAGGAAAGCAAGGGTG 664  
Db |||||||  
QY 605 AATGTTAATGGCATCGGCTTCCATTCGTTTCAAGACATCTAGGAAAGCAAGGGTG 664  
Db |||||||  
QY 665 GAAATGAGATTCAGATGATCGAAGTTTCCCATTCGAGTTCAATGGTGCAACATTC 724  
Db |||||||  
QY 665 GAAATGAGATTCAGATGATCGAAGTTTCCCATTCGAGTTCAATGGTGCAACATTC 724  
Db |||||||  
QY 725 ACATTCATGATGATCTCAATCTTGGGGGTAGGCGTAACAGCATAGGAAGCT 784  
Db |||||||  
QY 725 ACATTCATGATGATCTCAATCTTGGGGGTAGGCGTAACAGCATAGGAAGCT 784  
Db |||||||  
QY 785 GGCGGCGCCACCTTCACTAGAGCTCAAGAAATATTACAAATGAATTAAGAGTTTAGAA 844  
Db |||||||  
QY 785 GGCGGCGCCACCTTCACTAGAGCTCAAGAAATATTACAAATGAATTAAGAGTTTAGAA 844  
Db |||||||  
QY 845 CTGGCAGCGGATTCCTTTGCAAGCTATGATAGCTATATATCTCATGAAACTTG 904  
Db |||||||  
QY 845 CTGGCAGCGGATTCCTTTGCAAGCTATGATAGCTATATATCTCATGAAACTTG 904  
Db |||||||  
QY 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAAGATTTTACCTGCAAAATTTCTC 964  
Db |||||||  
QY 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAAGATTTTACCTGCAAAATTTCTC 964  
Db |||||||  
QY 965 CTTTTTGTGCTAGCAGGTTATTAGGTTCTCAGATAGATGATTCATATATGCTGCTATG 1024  
Db |||||||  
QY 965 CTTTTTGTGCTAGCAGGTTATTAGGTTCTCAGATAGATGATTCATATATGCTGCTATG 1024  
Db |||||||  
QY 1025 ABAACATTTGATGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084  
Db |||||||  
QY 1025 ABAACATTTGATGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084  
Db |||||||  
QY 1085 AAAAA 1089  
Db |||||||  
QY 1085 AAAAA 1089  
Db |||||||

## RESULT 2

TSP271917  
LOCUS TSP271917 1083 bp mRNA linear PLN 16-DEC-2000  
DEFINITION Triticum sp. mRNA for DP protein (dp gene).  
ACCESSION AJ271917  
VERSION AJ271917.1 GI:11877790  
KEYWORDS DP gene; E2F dimerization partner.  
SOURCE Triticum sp.

## ORGANISM

Triticum sp.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

## REFERENCE

1 Ramirez-Parra, E. and Gutierrez, C.  
Characterization of wheat DP, a heterodimerization partner of the  
Plant E2F transcription factor which stimulates E2F-DNA binding  
FEBS Lett. 486 (1), 73-78 (2000)

## MEDLINE

20562804

## PUBMED

11108846

## REFERENCE

2 (bases 1 to 1083)

## AUTHORS

Gutierrez, C.

## TITLE

Direct Submission

## JOURNAL

Submitted (03-FEB-2000) Gutierrez C., C.S.I.C.-U.A.M., Centro de  
Biologia Molecular, Cantoblanco, Madrid 28049, SPAIN  
Location/Qualifiers  
1. .1083  
/organism="Triticum sp."  
/mol\_type="mRNA"  
/db\_xref="taxon:4569"

## gene

1. .1083

## 5' UTR

<1. .13

## CDS

14. .799

/gene="dp"

/function="cell cycle E2F dimerization partner"

/codon\_start=1

/product="DP protein"

/protein\_id="Cac19034.1"

/db\_xref="GI:11877791"

/db\_xref="GOA:Q9FET1"

/db\_xref="UniProt/TREMBL:Q9FET1"

/translation="MAPRGGAAATAALDLTGWHILEASSVPPIPERGNAVQRKG  
AVDPKDKKERRVDAFNVLIALVIAKEKEIRWGLSNRYEKIKLEEVKELV  
NKRKKALLQIEKQFDLQNLKRNQTLSSAENVNGIRLPFVLVKTSRKARVEIE  
ISDDSKPAHFENFGAFTLHDLDSILEGVRGNSIGKAGRAFLH"

801. .1033

/gene="dp"

## ORIGIN

Query Match 59.5%; Score 648; DB 8; Length 1083;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 798; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 GGCGGTAAATGCGGTCCAAAGGAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGGAG 184

Db |||||||

QY 119 GGCGGTAAATGCGGTCCAAAGGAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGGAG 178

Db |||||||

QY 185 AAGCTGCGGACCGAGGATCACCGTTGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244

Db |||||||

QY 179 AAGCTGCGGACCGAGGATCACCGTTGGGGCTCCGGAGTACAGCAAAATAGTTTGT 238

Db |||||||

QY 245 GAGAAAGTTGAAGCAGGAGGAGCAACATACAAATGAGGTTGCAGAGAAATTTATTCA 304

Db |||||||

QY 239 GAGAAAGTTGAAGCAGGAGGAGCAACATACAAATGAGGTTGCAGAGAAATTTATTCA 298

Db |||||||

QY 305 GAGCTGAAGTCCATGCGACATATTTGTTCAAGGGTTTGTGAGAGAAATATTAGCGGAGA 364

Db |||||||

QY 299 GAGCTGAAGTCCATGCGACATATTTGTTCAAGGGTTTGTGAGAGAAATATTAGCGGAGA 358

Db |||||||

QY 365 GTGTATGATGCTTTCACGTTCTCATTGACATTCGTTTATTCGAAAGAAAGAGAG 424

Db |||||||

QY 359 GTGTATGATGCTTTCACGTTCTCATTGACATTCGTTTATTCGAAAGAAAGAGAG 418

Db |||||||

QY 425 ATACGCTGGATGGGCTTTCAAAATACAGATATGAAAAATAAAGAGCTTCAGGAAGTT 484

Db |||||||

QY 419 ATACGCTGGATGGGCTTTCAAAATACAGATATGAAAAATAAAGAGCTTCAGGAAGTT 478

Db |||||||

QY 485 CGTAAAGAACTCGTCAACAGATTAGGAAAGAGGCACTCTCCAGGAAATCGAAAAA 544

Db |||||||



```
Db 479 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 538
QY 545 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604
Db 539 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 598
QY 605 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGGAAGCAAGGGTG 664
Db 599 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGGAAGCAAGGGTG 658
QY 665 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGCCCAATTTTCAGTTCAATGGTGCAACCATTC 724
Db 659 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGCCCAATTTTCAGTTCAATGGTGCAACCATTC 718
QY 725 ACATTCGATGATGATCTCAATTCCTTTCAGGCGGTGAAGCGTAAACAGCATAGGAAGAGCT 784
Db 719 ACATTCGATGATGATCTCAATTCCTTTCAGGCGGTGAAGCGTAAACAGCATAGGAAGAGCT 778
QY 785 GCGCGCGCCACCTTCCTAGAGACTCAAGATATTTACAAATGAATTAAGTTGTAGAA 844
Db 779 GCGCGCGCCACCTTCCTAGAGACTCAAGATATTTACAAATGAATTAAGTTGTAGAA 838
QY 845 CTGGCACAGCGGATCTTTTTCACAGCTATGTATAGCTATATCTCATGAAAACTTG 904
Db 839 CTGGCACAGCGGATCTTTTTCACAGCTATGTATAGCTATATCTCATGAAAACTTG 898
QY 905 ACCTAGTTTATAGGACAGTCT 925
Db 899 ACCTAGTTTATAGGACAGTCT 919

RESULT 3
AX100712
LOCUS AX100712 168 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 9 from Patent WO0121644.
ACCESSION AX100712
VERSION AX100712.1 GI:13619660
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 9 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source Location/Qualifiers
1..168
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
1..>168
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36470.1"
/db_xref="GI:13619661"
/translation="MAHIGQGFDEKNIRRVYDAFNVLIALRVIAKEKKEIRWMGLSN
YRYEKIKLLEEV"

CDS
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGACTTCCTGTTTATTGCAAAAAAGAGATACGGTGGATG 436
Db 61 TTCAACGTTCTCATTCGACTTCCTGTTTATTGCAAAAAAGAGATACGGTGGATG 120

ORIGIN
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 544
Db 1 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 60
QY 545 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604
Db 61 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 120
QY 605 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGG 652
Db 121 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGG 168

RESULT 5
AX100710
LOCUS AX100710 153 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 7 from Patent WO0121644.
ACCESSION AX100710
VERSION AX100710.1 GI:13619658
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 7 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source Location/Qualifiers
1..153
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
1..>153
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36470.1"
/db_xref="GI:13619661"
/translation="MAHIGQGFDEKNIRRVYDAFNVLIALRVIAKEKKEIRWMGLSN
YRYEKIKLLEEV"
```

```
QY 437 GGCCTTTCAATTCAGATATGAAAAATAAAGAGCTTGAGGAAGTT 484
Db 121 GGCCTTTCAATTCAGATATGAAAAATAAAGAGCTTGAGGAAGTT 168

RESULT 4
AX100714
LOCUS AX100714 168 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 11 from Patent WO0121644.
ACCESSION AX100714
VERSION AX100714.1 GI:13619662
KEYWORDS Triticum monococcum
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 11 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source Location/Qualifiers
1..168
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
1..>168
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36471.1"
/db_xref="GI:13619663"
/translation="RKELVNKIRNKALLQEIEKQFDDLIQNIKLNRQTLSSAENVNG
IRLPFVLVKTSE"

ORIGIN
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 544
Db 1 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 60
QY 545 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604
Db 61 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 120
QY 605 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGG 652
Db 121 AATGTTAATGGCATCGCCTTCATTCGTTATGGTCAAGACATCTAGG 168

RESULT 5
AX100710
LOCUS AX100710 153 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 7 from Patent WO0121644.
ACCESSION AX100710
VERSION AX100710.1 GI:13619658
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 7 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source Location/Qualifiers
1..153
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
1..>153
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36470.1"
/db_xref="GI:13619661"
/translation="MAHIGQGFDEKNIRRVYDAFNVLIALRVIAKEKKEIRWMGLSN
YRYEKIKLLEEV"
```

```
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<1..>153
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36469.1"
/db_xref="GI:13619659"
/translation="DKDRKKEKAAAPRIITGWGLREYSKIVCEKVEAKGRTTYNEVADE
IYSELKS"

ORIGIN
Query Match 14.0%; Score 153; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 6.4e-71; Indels 0; Gaps 0;
Matches 153; Conservative 0; Mismatches 0;

QY 164 GATAAAGATAGGAAGAGGAGGCTGGCGCACCGAGGATCACCGGTTGGGGGCTCCGC 223
Db 1 GATAAAGATAGGAAGAGGAGGCTGGCGCACCGAGGATCACCGGTTGGGGGCTCCGC 60
QY 224 GAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGAG 283
Db 61 GAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGAG 120
QY 284 GTTGCAGACCAAAATTTATTTCAGAGCTGAAGTCC 316
Db 121 GTTGCAGACCAAAATTTATTTCAGAGCTGAAGTCC 153

RESULT 6
AX100716 150 bp DNA linear PAT 10-APR-2001
LOCUS
DEFINITION
Sequence 13 from Patent WO0121644.
ACCESSION
AX100716 GI:13619664
VERSION
KEYWORDS
SOURCE
Triticum monococcum
ORGANISM
Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
Wheat dp proteins and uses thereof
Patent: WO 0121644-A 5 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source
Location/Qualifiers
1..150
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<1..>156
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36468.1"
/db_xref="GI:13619657"
/translation="ARAAVAPRGGAATAALDLTGWHILEASSVPLPERGGNAV
QRKGAADP"

ORIGIN
Query Match 10.5%; Score 114; DB 6; Length 156;
Best Local Similarity 100.0%; Pred. No. 7e-50; Indels 0; Gaps 0;
Matches 114; Conservative 0; Mismatches 0;

QY 8 GCACGAGCGCAATGGCGCTCCCGCGCGAGAGTGCTGGCGCGCTACCGCGGCACTG 67
Db 1 GCACGAGCGCAATGGCGCTCCCGCGCGAGAGTGCTGGCGCGCTACCGCGGCACTG 60
QY 68 GACCTGACCGCGGTGCACATTCGAGCTCCAGTGTCCTCCCGCTCCCGAA 121
Db 61 GACCTGACCGCGGTGCACATTCGAGCTCCAGTGTCCTCCCGCTCCCGAA 114

RESULT 8
SLU39059 1132 bp mRNA linear PLN 18-NOV-1996
LOCUS
DEFINITION
Lycopersicon esculentum myo-inositol monophosphatase 3 mRNA,
complete cds.
ACCESSION
U39059
VERSION
U39059.1 GI:1098970
KEYWORDS
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Plant inositol monophosphatase is a lithium-sensitive enzyme
encoded by a multigene family
Plant Cell 7 (12), 2175-2185 (1995)
JOURNAL
MEDLINE
PUBMED
96351935
8718627
REFERENCE
2 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Direct Submission
Submitted (20-OCT-1995) Glenda E. Gillaspay, Plant Biology,
University of California, 111 Koshland Hall, Berkeley, CA 94720,

CDS
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<1..>150
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36472.1"
/db_xref="GI:13619665"
/translation="KARVEIEISDDSKFAHFPENGAPFTLHDDLILLEGVRRNSIGRA
GRATLH"

ORIGIN
Query Match 13.8%; Score 150; DB 6; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-69; Indels 0; Gaps 0;
Matches 150; Conservative 0; Mismatches 0;

QY 653 AAAGCAAGGTGGAATTCAGATTCAGATGACTCGAAGTTGCCCATTTCCGATTTCAAT 712
Db 1 AAAGCAAGGTGGAATTCAGATTCAGATGACTCGAAGTTGCCCATTTCCGATTTCAAT 60
QY 713 GTTGCCACCATTCATGATGATGATCTCAATCTTGAGGGGTAAAGCGGTACACG 772
Db 61 GTTGCCACCATTCATGATGATGATCTCAATCTTGAGGGGTAAAGCGGTAAACG 120
QY 773 ATAGGAAGAGCTGGCGCGCCACCCCTTCAC 802
```

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gathersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hghri.nih.gov](mailto:nisc_mgc@hghri.nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Mastello, C., Maskeri, B., Macrián, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Scantrippo, S., Thomas, P.J., Touchman, J.W.,  
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IPAK Plate: 62 Row: k Column: 19  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 6005885.

	FEATURES	SOURCE
1	100% Pure Cotton	100%
2	100% Pure Cotton	100%
3	100% Pure Cotton	100%
4	100% Pure Cotton	100%
5	100% Pure Cotton	100%
6	100% Pure Cotton	100%
7	100% Pure Cotton	100%
8	100% Pure Cotton	100%
9	100% Pure Cotton	100%
10	100% Pure Cotton	100%
11	100% Pure Cotton	100%
12	100% Pure Cotton	100%
13	100% Pure Cotton	100%
14	100% Pure Cotton	100%
15	100% Pure Cotton	100%
16	100% Pure Cotton	100%
17	100% Pure Cotton	100%
18	100% Pure Cotton	100%
19	100% Pure Cotton	100%
20	100% Pure Cotton	100%
21	100% Pure Cotton	100%
22	100% Pure Cotton	100%
23	100% Pure Cotton	100%
24	100% Pure Cotton	100%
25	100% Pure Cotton	100%
26	100% Pure Cotton	100%
27	100% Pure Cotton	100%
28	100% Pure Cotton	100%
29	100% Pure Cotton	100%
30	100% Pure Cotton	100%
31	100% Pure Cotton	100%
32	100% Pure Cotton	100%
33	100% Pure Cotton	100%
34	100% Pure Cotton	100%
35	100% Pure Cotton	100%
36	100% Pure Cotton	100%
37	100% Pure Cotton	100%
38	100% Pure Cotton	100%
39	100% Pure Cotton	100%
40	100% Pure Cotton	100%
41	100% Pure Cotton	100%
42	100% Pure Cotton	100%
43	100% Pure Cotton	100%
44	100% Pure Cotton	100%
45	100% Pure Cotton	100%
46	100% Pure Cotton	100%
47	100% Pure Cotton	100%
48	100% Pure Cotton	100%
49	100% Pure Cotton	100%
50	100% Pure Cotton	100%
51	100% Pure Cotton	100%
52	100% Pure Cotton	100%
53	100% Pure Cotton	100%
54	100% Pure Cotton	100%
55	100% Pure Cotton	100%
56	100% Pure Cotton	100%
57	100% Pure Cotton	100%
58	100% Pure Cotton	100%
59	100% Pure Cotton	100%
60	100% Pure Cotton	100%
61	100% Pure Cotton	100%
62	100% Pure Cotton	100%
63	100% Pure Cotton	100%
64	100% Pure Cotton	100%
65	100% Pure Cotton	100%
66	100% Pure Cotton	100%
67	100% Pure Cotton	100%
68	100% Pure Cotton	100%
69	100% Pure Cotton	100%
70	100% Pure Cotton	100%
71	100% Pure Cotton	100%
72	100% Pure Cotton	100%
73	100% Pure Cotton	100%
74	100% Pure Cotton	100%
75	100% Pure Cotton	100%
76	100% Pure Cotton	100%
77	100% Pure Cotton	100%
78	100% Pure Cotton	100%
79	100% Pure Cotton	100%
80	100% Pure Cotton	100%
81	100% Pure Cotton	100%
82	100% Pure Cotton	100%
83	100% Pure Cotton	100%
84	100% Pure Cotton	100%
85	100% Pure Cotton	100%
86	100% Pure Cotton	100%
87	100% Pure Cotton	100%
88	100% Pure Cotton	100%
89	100% Pure Cotton	100%
90	100% Pure Cotton	100%
91	100% Pure Cotton	100%
92	100% Pure Cotton	100%
93	100% Pure Cotton	100%
94	100% Pure Cotton	100%
95	100% Pure Cotton	100%
96	100% Pure Cotton	100%
97	100% Pure Cotton	100%
98	100% Pure Cotton	100%
99	100% Pure Cotton	100%
100	100% Pure Cotton	100%

gene

CDS

## ORIGIN

[illegible]

## RESULT 10

CR628367

LOCUS	CR628367	246093 bp	DNA linear	HTG 28-SEP-2000
DEFINITION	Mus musculus chromosome X clone RP23-75B12, WORKING DRAFT SEQUENCE, 2 unordered pieces.			

```

FEATURES             USA
source
    Location/Qualifiers
      1..1132
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /strain="VEXT Cherry"
        /db_xref="taxon:4081"
      13..819
        /codon_start=1
        /product="myo-inositol monophosphatase 3"
        /protein_id="AAB19031.1"
        /db_xref="GI:1098971"
        /translation="MAQNSGVGEQFLVAVEAAKAGEIIRGEFYKTKVHEHKGMVDLV
        TETDKACDGFIFNHLKQRFPSHKFIGEETAAACGNFELTDEPTWIVDPDLTKVHEHKGMVDLV
        FPFVCVSIIGLTIEGTEGQFVWVNPYIIDEFLTGDGKGAFLNGKPIKVSQSOELVALL
        ATEAGNDRKLVVDATGRIINSLFKVSRILCMGSCGALNCVACGRLLDLFYELEFGG
        PNDVAGGAVIVKEAGGVEFDPGSGSEFDITARRVAAATNAHDKAFTKAIKATNR"
      CDS

```

## ORIGIN

[illegible]

RESULT 9				
BC028028	EC028028	1630 bp	mRNA	linear PRI 30-SEP-2003
LOCUS				
DEFINITION	Homo sapiens syntaxin binding protein 3, mRNA (cDNA clone IMAGE:5219755), partial cds.			
ACCESSION	BC028028			
VERSION	BC028028.1	GI:20380108		

REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

```

COMMENT
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 28, 2004 this sequence version replaced gi:51966191.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm75B12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 245844 bases at least Q40
Consensus quality: 245914 bases at least Q30
Consensus quality: 245977 bases at least Q20
Insert size: 245993; sum-of-contigs
Insert size: 220507; 3.1% error; agarose-fp
Quality coverage: 10.26x in Q20 bases; sum-of-contigs Quality
coverage: 12.01x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 132816: contig of 132816 bp in length
* 132817 132916: gap of 100 bp
* 132917 246093: contig of 113177 bp in length.
FEATURES             source
    source
        1..246093
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="X"
            /clone="RP23-75B12"
            /clone_lib="RPCI-23"
        misc_feature
            1..132816
                /notes="assembly fragment:04407
                fragment chain:1"
            132917..246093
                /notes="assembly fragment:04662
                fragment chain:1
                clone_end:SP6
                vector_side:right"
ORIGIN
Query Match          5.2%; Score 57; DB 2; Length 246093;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1033 GATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
|
|
|
Db 180762 GATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180818
|
|
|
RESULT 11
BC002750
LOCUS
DEFINITION Homo sapiens chromosome 11 open reading frame 10, mRNA (cdna clone
MGC:3384 IMAGE:3633279), complete cds.
ACCESSION BC002750
VERSION BC002750.2 GI:38014315
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 444)  
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. MGC clone  
 generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
 2 (bases 1 to 444)  
 Strausberg,R.  
 Direct Submission  
 Submitted (05-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 On Oct 28, 2003 this sequence version replaced gi:12803818.  
 Contact: MGC help desk  
 Email: cgapps-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: http://www.nisc.nih.gov/  
 Contact: nisc.mgc@nih.gov  
 Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,  
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,  
 Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,  
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,  
 Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAL Plate: 12 Row: k Column: 24  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 7656933.

FEATURES  
source

1..444  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue="MGC:3384 IMAGE:3633279"  
 /tissue\_type="Uterus, endometrium adenocarcinoma"  
 /clone\_lib="NIH MGC 44"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..444  
 /gene="Clorf10"  
 /db\_xref="LocusID:746"  
 27..266  
 /gene="Clorf10"  
 /codon\_start=1

## gene

## CDS

/product="chromosome 11 open reading frame 10"

/protein\_id="AAH02750.1"

/db\_xref="GI:12803819"

/db\_xref="LocusID:746"

/translation="MELEAWSRYTSPNPAVPHLTVLLAIGMPTTAWFFVYEVST  
KYTRDIYKELLISLVASLPMFGVLFLLWGIYV"

#### ORIGIN

Query Match 5.1%; Score 56; DB 9; Length 444;

Best Local Similarity 100.0%; Pred. No. 1.3e-18;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

||||| 1089

Db 377 ATAGCAA 432

||||| 432

#### RESULT 12

BC015791

LOCUS

BC015791 Homo sapiens, clone IMAGE:4866083, mRNA. PRI 11-OCT-2001

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

BC Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hailao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 34 Row: j Column: 17

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

#### FEATURES

source

1. .1980

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4866083"

/tissue="Colon, adenocarcinoma"

/clone\_lib="NIH MGC 15"

/lab\_host="DH10B-R"

/note="Vector: pOT87"

#### ORIGIN

Query Match

Best Local Similarity

5.1%; Score 56; DB 9; Length 1980;

100.0%; Pred. No. 1.2e-18;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

||||| 1089

Db 1914 ATAGCAA 1969

||||| 1969

#### RESULT 13

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

||||| 1089

Db 1914 ATAGCAA 1969

||||| 1969

#### RESULT 13

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

||||| 1089

Db 1914 ATAGCAA 1969

||||| 1969

#### RESULT 13

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

AY118692

LOCUS

AY118692

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE



Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Nasillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 139 Row: i Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES

## Location/Qualifiers

## Source

1..3430

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5732786"

/tissue\_type="Brain, hippocampus"

/clone\_lib="NIH\_MGC\_124"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

&lt;1..3430

/gene="KIAA0602"

/db\_xref="LocusID:23241"

&lt;1..2787

/gene="KIAA0602"

/codon\_start=1

/product="KIAA0602 protein"

/protein\_id="AAH65220.1"

/db\_xref="GI:40850908"

/db\_xref="LocusID:23241"

/translation="PPGSHVTAPPPSARPPRVGRGPAAPRRGGGAGAMAERG  
RIGLPGAGALNTPVPMNLFATVEVDGSPSCVPRLCSTLTKLVVFKLEKELISV  
IAVQGGKRIILRSHEIILPPSQVETDLALTSLQYPHFLKREGNKQIMLQRRKRY  
KNRTILGYKTLAGSTMAEVMQHPSEGGVLSLCSSIKEAPVKAEEIWIASLSOPI  
DHEDTMOAGPKAKSTDNYSEIEYSEFSSEQASDDAVQGDLDDEDDVDGPKKORR  
SIVRTSMTRQNFQKQVALLRRFKVSDVLDSEQDPAEHIPEAEEDLDLLYDTLDM  
EHPSSGPMDEDDSVLSTPKPLRPYFGLSHSSQTEIGSIHARSHEKPPSPADY  
PEKTRSLGRQPSDSVDTVALGVPGRHPGQPEDPEAEASTLDVFTERLPPSGRI  
TKTESLIVIPSTRSEGRGRSTSLKEROAQRQNERANSIDNERCPDARSQIQIP  
RKTVDQLNHLISDDLPENILVNTSDQGFSLDVLORHTLPVCTCSADYQAA  
ESTIVSRIORYCNKSOPPTPVKIAVAGAHVLSAILRLFEQLSHKTPDWLGYMREL  
VIPGSHFVARVILGSDVRYNNPFQDLARLDLENKLEAGSAVQDTPDIVSRITQYIAG  
ANCAHLPFAEAWLITYKQSPDEESSQKFIFFGVVKGIVPESSATSGSDDAAPSG  
SGTLSTPPSPAPAAKEASPTPPSPSVSGSLSPSQGVGAEIMGLOVDYNTAAQPAD  
RRDAEKDLPVTNKLTPRSLQVSRLLPSSGEAAATPTMSMTVVTKEKNKVMFLP  
KKAQDKVESKQCTEGISGLICTARQQQNMRLVLIDGVECSDVKFFQLAAQWSSHVK  
HFPICIFGHSKATP"

## ORIGIN

Query Match 5.1%; Score 56; DB 9; Length 3430;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1034 ATAGCAA 1089  
|||  
Db 3316 ATAGCAAA 3371  
|||

Search completed: February 6, 2005, 20:16:32  
Job time : 5152 secs



***This Page Blank (usp10)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 15:46:16 ; Search time 226 Seconds  
(without alignments)  
7884.537 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgagcagcgcaaa.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCUS COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	5.1	491	US-09-311-021-191	Sequence 191, Appl
2	54	5.0	176	US-09-621-976-13903	Sequence 13903, A
3	54	5.0	630	US-08-185-414E-1	Sequence 1, Appl
4	54	5.0	985	US-09-322-409-25	Sequence 25, Appl
5	54	5.0	985	US-09-322-409-27	Sequence 27, Appl
6	54	5.0	985	US-09-451-527-25	Sequence 25, Appl
7	54	5.0	985	US-09-451-527-27	Sequence 27, Appl
8	54	5.0	1023	US-08-252-966B-16	Sequence 16, Appl
9	54	5.0	1050	US-09-482-273-58	Sequence 58, Appl
10	54	5.0	1358	US-09-949-016-463	Sequence 463, Appl
11	54	5.0	1459	US-09-537-654-3	Sequence 3, Appl
12	54	5.0	1810	US-09-800-729-73	Sequence 73, Appl
13	54	5.0	1811	US-09-800-729-77	Sequence 77, Appl
14	54	5.0	2406	US-09-594-506-37	Sequence 37, Appl
15	54	5.0	3438	US-10-164-595-29	Sequence 29, Appl
16	53	4.9	150	US-09-621-976-8656	Sequence 8656, Ap
17	53	4.9	165	US-09-621-976-8127	Sequence 8127, Ap
18	53	4.9	194	US-09-621-976-801	Sequence 801, Appl
19	53	4.9	233	US-09-621-976-10675	Sequence 10675, A
20	53	4.9	249	US-09-621-976-19144	Sequence 19144, A
21	53	4.9	674	US-09-620-405B-465	Sequence 465, Appl
22	53	4.9	674	US-09-433-826B-465	Sequence 465, Appl
23	53	4.9	674	US-09-604-287A-465	Sequence 465, Appl
24	53	4.9	674	US-09-834-759-465	Sequence 465, Appl
25	53	4.9	674	US-09-590-751A-465	Sequence 465, Appl
26	53	4.9	674	US-09-551-621-465	Sequence 465, Appl
27	53	4.9	708	US-09-270-767-13081	Sequence 13081, A

```

28 53 4.9 857 1 US-08-308-883-1 Sequence 1, Appli
29 53 4.9 857 1 US-08-730-163-1 Sequence 1, Appli
30 53 4.9 857 3 US-08-256-799-1 Sequence 1, Appli
31 53 4.9 857 3 US-08-462-437-1 Sequence 1, Appli
32 53 941 4 US-09-205-258-186 Sequence 186, App
33 53 4.9 1066 1 US-08-157-101A-4 Sequence 4, Appli
34 53 4.9 1181 3 US-09-149-476-310 Sequence 310, App
35 53 4.9 1196 4 US-09-065-040-2 Sequence 2, Appli
36 53 4.9 1212 3 US-09-149-476-186 Sequence 186, App
37 53 4.9 1307 2 US-08-960-022-17 Sequence 17, Appl
38 53 4.9 1359 3 US-09-387-574-11 Sequence 11, Appl
39 53 4.9 1359 3 US-09-668-096-11 Sequence 11, Appl
40 53 4.9 1525 4 US-09-461-325-110 Sequence 110, App
41 53 4.9 1525 4 US-10-012-542-110 Sequence 110, App
42 53 4.9 1525 4 US-10-115-123-110 Sequence 110, App
43 53 4.9 1534 1 US-08-300-903A-6 Sequence 6, Appli
44 53 4.9 1534 4 US-08-988-197-6 Sequence 6, Appli
45 53 4.9 1534 4 US-10-385-072-6 Sequence 6, Appli

```

#### ALIGNMENTS

```

RESULT 1
US-09-311-021-191
; Sequence 191, Application US/093111021
; Patent No. 6706869
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J. Inc.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6300-11A
; CURRENT APPLICATION NUMBER: US/09/311,021
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 191
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-311-021-191

```

```

Query Match 5.1%; Score 55; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
    |||
Db 430 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 484

```

```

RESULT 2
US-09-621-976-13903
; Sequence 13903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13903
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-13903

```

```
Query Match          5.0%; Score 54; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 85 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 138

RESULT 3
US-08-185-414E-1
; Sequence 1, Application US/08185414E
; Patent No. 5556953
; GENERAL INFORMATION:
; APPLICANT: Zhang, Lei
; APPLICANT: Vijay, Hari M.
; APPLICANT: Rode, Harold
; TITLE OF INVENTION: ALLERGEN OF CLADOSPORIUM HERBARUM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George A. Seaby
; ADDRESSEE: Seaby and Maclean
; STREET: 880 Wellington Street, Suite 708
; CITY: Ottawa
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,414E
; FILING DATE: January 24, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George A. Seaby
; REGISTRATION NUMBER: 24,034
; REFERENCE/DOCKET NUMBER: 1747
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 232-5815
; TELEFAX: (613) 232-5831
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA and its encoded protein
; FEATURE: Open reading frame of CH2.1
US-08-185-414E-1

Query Match          5.0%; Score 54; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 544 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 597

RESULT 4
US-09-322-409-25
; Sequence 25, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; MOLECULE TYPE: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1

Query Match          5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 65 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 5
US-09-322-409-27/c
; Sequence 27, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; MOLECULE TYPE: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-27

Query Match          5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 921 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 974

RESULT 6
US-09-451-527-25
; Sequence 25, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; MOLECULE TYPE: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
```

```

/ COUNTRY: USA
/ ZIP: 98101-2347
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/252,966B
/ FILING DATE: 01-JUN-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Shelton, Dennis K.
/ REGISTRATION NUMBER: 26,997.
/ REFERENCE/DOCKET NUMBER: FHCRL7694
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 682-8100
/ TELEFAX: (206) 224-0779
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1023 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to mRNA
/ DESCRIPTION: clone 20; see Figure 27
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Mus musculus
/ US-08-252-966B-16

Query Match 5.0%; Score 54; DB 1; Len:
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0;

QY 1036 ACACAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
DB 946 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

RESULT 9
US-09-482-273-58
/ Sequence 58, Application US/09482273
/ Patent No. 6534631
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 71 Human Secreted Proteins
/ FILE REFERENCE: PZ030P1
/ CURRENT APPLICATION NUMBER: US/09/482,273
/ CURRENT FILING DATE: 2000-01-13
/ EARLIER APPLICATION NUMBER: PCT/US99/15849
/ EARLIER FILING DATE: 1999-07-14
/ EARLIER APPLICATION NUMBER: 60/092,921
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,922
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,956
/ EARLIER FILING DATE: 1998-07-15
/ NUMBER OF SEQ ID NOS: 267
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 58
/ LENGTH: 1050
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-482-273-58

Query Match 5.0%; Score 54; DB 4; Len:
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0;

QY 1036 ACACAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
DB 997 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```



Tue Feb 8 10:33:59 2005

```
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Thiamine Biosynthetic Enzymes
; FILE REFERENCE: BB1372 US NA
; CURRENT APPLICATION NUMBER: US/09/594,506
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,556
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-594-506-37

Query Match          5.0%; Score 54; DB 4; Length 2406;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2351 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2404

RESULT 15
US-10-164-595-29
; Sequence 29, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (527)..(2701)
; OTHER INFORMATION:
US-10-164-595-29

Query Match          5.0%; Score 54; DB 4; Length 3438;
Best Local Similarity 100.0%; Pred. No. 2.6e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3345 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3398

Search completed: February 6, 2005, 17:39:37
Job time : 228 secs
```

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 17:09:27 ; Search time 4251 Seconds  
(without alignments)  
9751.115 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	9.8	694	9	CL902823
2	98	9.0	454	6	CD892664
3	88	8.1	649	6	CD882927
4	59	5.4	301	6	CA802821
5	59	5.4	318	1	AJ774826
6	59	5.3	157	6	CB371210
7	56	5.1	169	1	AL697897
8	56	5.1	177	4	BG179244
9	56	5.1	214	7	CO249876
10	56	5.1	223	6	CA802105
11	56	5.1	262	7	CN005830
12	56	5.1	269	2	AW734913
13	56	5.1	287	1	AL046990
14	56	5.1	290	4	BI756116
15	56	5.1	294	1	AA413331
16	56	5.1	310	7	CK593793
17	56	5.1	326	6	CB846427
18	56	5.1	335	7	CN840208
19	56	5.1	388	6	CA934974
20	56	5.1	395	4	BI337568
21	56	5.1	401	1	AV656973
22	56	5.1	443	5	BX474270
23	56	5.1	453	4	BI384539
24	56	5.1	455	1	AA638556

25	56	5.1	458	7	CV497760
26	56	5.1	460	7	CK085723
27	56	5.1	489	4	BI377034
28	56	5.1	491	6	CD722238
29	56	5.1	497	4	BG164383
30	56	5.1	525	7	CK938853
31	56	5.1	567	4	BI377687
32	56	5.1	567	6	C92906
33	56	5.1	586	2	BE907643
34	56	5.1	611	7	CN992367
35	56	5.1	628	4	BI861964
36	56	5.1	689	2	BE394168
37	56	5.1	806	5	BU958336
38	56	5.1	951	2	BE879000
39	55	5.1	113	2	BE229946
40	55	5.1	120	4	BI942947
41	55	5.1	141	7	CF317851
42	55	5.1	147	2	AW481765
43	55	5.1	149	7	CK121803
44	55	5.1	159	7	CO168353
45	55	5.1	170	2	BE229931

## ALIGNMENTS

RESULT 1  
CL902823  
LOCUS  
DEFINITION  
CSHC2198 1188HC library Triticum aestivum genomic clone  
1188HC01N01, genomic survey sequence.  
CL902823  
CL902823.1 GI:51666086  
GSS.  
Triticum aestivum (bread wheat)  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 694)  
Lamoureux,D., Peterson,D.G., Li,W., Fellers,J.P. and Gill,B.S.  
Cot-based cloning and sequencing (CBCS) efficiently removes  
sequence repeats and increases gene ratio in bread wheat  
Unpublished (2004)  
Contact: Gill BS  
Department of Plant Pathology  
Kansas State University  
4024 Throckmorton, Manhattan, KS 66506-5502, USA  
Tel: 785 532 1391  
Fax: 785 532 5692  
Email: bagill@ksu.edu  
Seq primer: T7  
Class: sheared ends.  
Location/Qualifiers  
1..694  
/organism="Triticum aestivum"  
/mol\_type="genomic DNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="1188HC01N01"  
/tissue\_type="whole plant"  
/dev\_stage="young shoot"  
/clone\_lib="1188HC library"

FEATURES  
source

Query Match 9.8%; Score 107; DB 9; Length 694;  
Best Local Similarity 100.0%; Pred. No. 2.6e-40;  
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 15 CCGCATGGCGCTCCCGCGCGAGCTGCTGGCGCGCTACCGCGCACTGGACTGA 74  
|||||  
Db 549 CCGCAATGGCGCTCCCGCGCGAGCTGCTGGCGCGCTACCGCGCACTGGACTGA 608  
|||||

```

QY 75 CCGCGTGCACATTCCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGAA 121
    |||||||
Db 609 CCGCGTGCACATTCCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGAA 655

RESULT 2
CD892664
LOCUS G118.121J23F010725 G118 Triticum aestivum cDNA clone G118121J23,
DEFINITION mRNA sequence.
ACCESSION CD892664
VERSION CD892664.1 GI:32663198
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 454)
Genoplante, a major partnership french program in plant genomics
Genoplante, unpublished (2003)
JOURNAL Contact: Genoplante
COMMENT Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
    source
    1..454
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="G1181J23"
    /tissue_type="grain (118 degrees per day after
    pollination)"
    /clone_lib="G118"

ORIGIN
    Query Match 9.0%; Score 98; DB 6; Length 454;
    Best Local Similarity 99.3%; Pred. No. 5.9e-36;
    Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 689 AAGTTTGGCCATTCGAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 748
    |||||||
Db 92 AAGTTTGGCCATTCGAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 151

QY 749 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 808
    |||||||
Db 152 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 211

QY 809 CTCAGATATTTACAAATGATTTAAAGT 837
    |||||||
Db 212 CTCAGATATTTACAAATGATTTAAAGT 240

RESULT 3
CD882927
LOCUS F1.111L02F010430 F1 Triticum aestivum cDNA clone F111L02, mRNA
DEFINITION sequence.
ACCESSION CD882927
VERSION CD882927.1 GI:32645392
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 649)
Genoplante, a major partnership french program in plant genomics
Genoplante, unpublished (2003)
JOURNAL Contact: Genoplante
COMMENT Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
    source
    1..454
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="G1181J23"
    /tissue_type="grain (118 degrees per day after
    pollination)"
    /clone_lib="G118"

ORIGIN
    Query Match 9.0%; Score 98; DB 6; Length 454;
    Best Local Similarity 99.3%; Pred. No. 5.9e-36;
    Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 689 AAGTTTGGCCATTCGAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 748
    |||||||
Db 92 AAGTTTGGCCATTCGAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 151

QY 749 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 808
    |||||||
Db 152 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 211

QY 809 CTCAGATATTTACAAATGATTTAAAGT 837
    |||||||
Db 212 CTCAGATATTTACAAATGATTTAAAGT 240

RESULT 4
CA802821
LOCUS sau42g02.y1 Gm-c1071 Glycine max cDNA clone soybean EST 01-JUL-2004
DEFINITION Gm-c1071-3748 5', mRNA sequence.
ACCESSION CA802821
VERSION CA802821.1 GI:26059907
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 301)
Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swallow, T., Gibbons, M., Pape, D., Harvey, N.,
Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Putative full length read vector to vector length is 467
Seq primer: -40SP from Gibco
High quality sequence stop: 286.
Location/Qualifiers

FEATURES
    source
    1..649
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="F111L02"
    /tissue_type="leaf one"
    /clone_lib="F1"

ORIGIN
    Query Match 8.1%; Score 88; DB 6; Length 649;
    Best Local Similarity 100.0%; Pred. No. 4e-31;
    Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 GCAAGAGAGAGAGAGAGATCGGTGGATGGCGCTTCAAAATACAGATATGAAAAATA 466
    |||||||
Db 119 GCAAGAGAGAGAGAGAGATCGGTGGATGGCGCTTCAAAATACAGATATGAAAAATA 178

QY 467 AAGAAGCTTGAAGGAGTTCGTAAAGAAC 494
    |||||||
Db 179 AAGAAGCTTGAAGGAGTTCGTAAAGAAC 206

RESULT 4
CA802821
LOCUS sau42g02.y1 Gm-c1071 Glycine max cDNA clone soybean EST 01-JUL-2004
DEFINITION Gm-c1071-3748 5', mRNA sequence.
ACCESSION CA802821
VERSION CA802821.1 GI:26059907
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 301)
Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swallow, T., Gibbons, M., Pape, D., Harvey, N.,
Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Putative full length read vector to vector length is 467
Seq primer: -40SP from Gibco
High quality sequence stop: 286.
Location/Qualifiers

FEATURES
    source
    1..649
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="F111L02"
    /tissue_type="leaf one"
    /clone_lib="F1"

ORIGIN
    Query Match 8.1%; Score 88; DB 6; Length 649;
    Best Local Similarity 100.0%; Pred. No. 4e-31;
    Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 GCAAGAGAGAGAGAGATCGGTGGATGGCGCTTCAAAATACAGATATGAAAAATA 466
    |||||||
Db 119 GCAAGAGAGAGAGAGATCGGTGGATGGCGCTTCAAAATACAGATATGAAAAATA 178

QY 467 AAGAAGCTTGAAGGAGTTCGTAAAGAAC 494
    |||||||
Db 179 AAGAAGCTTGAAGGAGTTCGTAAAGAAC 206

```

	Matches	59;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1031	TTGATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA linear EST 18-MAR-2003 TGSTZYG9C09.xl TgME49 3 day invitro bradyzoite Toxoplasma gondii cDNA clone TGSTZYg9C09.xl 3' mRNA sequence.								1089
Dd	228	TTGATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA								286
RESULT 6										
CBS71210	157 bp mRNA									
DEFINITION										
ACCESSION										
VERSION										
KEYWORDS										
SOURCE										
ORGANISM										
REFERENCE										
AUTHORS	Tang,K., Cole,R.E., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M., Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons.M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili.R., Ronko,I., Kennedy,S., Meguire,U., Waterston,R. and Wilson,R.  Toxoplasma EST Project Unpublished (2001) Contact: Clifton, S.  Toxoplasma EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: toxo@watson.wustl.edu  Contact David Sibley (toxoeat@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.  Putative full length read vector to vector length is Seq primer: -40UP from GIBCO.									
TITLE										
JOURNAL										
COMMENT										
FEATURES										
SOURCE	1..157 /organism=Toxoplasma gondii" /mol_type="mRNA" /db_xref="taxon:5811" /clone=TGSTZYg9C09.xl" /dev_stage="3 day" /lab_host="ElectronTen Blue cells" /clone_lib=TgME49 3 day Invitro bradyzoite" /note=Vector: pBluescript SK; Site_1: NotI; Site_2: Smal; cDNA was reverse transcribed using Superscript II and an anchored (?wobble') oligo dT primer, containing a NOTI restriction site, according to standard protocols. Blunt second strand cDNA was digested with NOTI restriction endonuclease, purified and size fractionated with a SizeSep 400 spun column (Amersham) and directionally cloned into the Smal and NotI sites of pBluescript SK. Following electroporation into ElectronTen Blue cells (Stratagene), 5 x 10^6 primary CFU were bottle amplified in semisolid culture media [2xLB+15% SeaPrep agarosed (BMA)]. After 48 hrs at 30 degrees Celsius, cells were harvested, pooled and stored in 2xLB + 15% glycerol."									
ORIGIN										
Query Match	5.3%; Score 58; DB 6; Length 157;									
Best Local Similarity	100.0%; Fred.No. 1.4e-16;									
Matches	58;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1032	TGATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA								1089
Dd	55	TGATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA								112
RESULT 7										

AL697897	AL697897	169 bp	mrna	linear	EST 04-SEP-2003
LOCUS	DKFZp686A19105.r1.686 (synonym: hlcc3)			Homo sapiens	cdna clone
DEFINITION	DKFZp686A19105.5', mRNA sequence.				
ACCESSION	AL697897				
VERSION	AL697897.1	GI:19618437			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 169)				
TITLE	Koehler, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.				
JOURNAL	EST (Koeher, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)				
COMMENT	Unpublished (1999)				
	Contact: MIPS				
	MIPS				
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany				
	This is the 5' sequence of the clone insert				
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;				
	sequenced by BMF (Biomedical Research Center at the Charite, Berlin/Germany) within the cdna sequencing consortium of the German Genome Project.				
	No BL sequence available.				
	This clone (DKFZp686A19105) is available at the RZPD in Berlin.				
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.				

```

source
1. .169
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686A19105"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/notes="Vector: pTriplex2; Site_1: SfIIA; Site_2: SfIIB;
cDNA-collection"

ORIGIN
Query Match          5.1%; Score 56; DB 1; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1034 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
        |||
        60 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 115

RESULT 8
BG179244
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 177)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: DCID/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

```

RESULT 9	
CO249876	214 bp mRNA linear EST 23-JUN-2004
LOCUS	
DEFINITION	AGENCOURT 26529183 NIH MGC 212 Homo sapiens cDNA clone
ACCESSION	IMAGE:30925067 5', mRNA sequence.
VERSION	CO249876
KEYWORDS	CO249876.1 GI:49121395
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1 (bases 1 to 214)
COMMENT	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: <a href="mailto:cgabbs-re@mail.nih.gov">cgabbs-re@mail.nih.gov</a> Tissue Procurement: Mary Hendrix cDNA Library preparation: M. Bento Soares, University of Iowa DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: NDAM1168 row: i column: 12 High quality sequence stop: 214. Location/Qualifiers 1. .214 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30925067" /tissue_type="Chondroarcoma Lung Metastasis cell lines" /lab_host="DH10B (T1 phage resistant)" /clone_lib="NIH_MGC_212" /note="Organ: Lung; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose
FEATURES	
source	



```

RESULT 12
AW734913
LOCUS
DEFINITION AW734913 269 bp mRNA linear EST 14-JUL-2004
Gm-c1016-10054 5', mRNA sequence.
ACCESSION
VERSION AW734913
KEYWORDS
SOURCE EST.
ORGANISM Glycine max (soybean)
Glycine max
Glycine max
REFERENCE 1 (bases 1 to 269)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Public Soybean EST Project
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 227.
FEATURES
source
1..269
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-10054"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/clone_lib="Gm-c1016"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelting."
ORIGIN
Query Match 5.1%; Score 56; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1034 ATACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 207 ATACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 262
RESULT 13
AL046990
LOCUS
DEFINITION AL046990 287 bp mRNA linear EST 04-SEP-2003
DXFZ586L2017 r1.586 (synonym: hutel) Homo sapiens cDNA clone
DXFZ586L2017 5', mRNA sequence.
ACCESSION
VERSION AL046990
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Homo sapiens
REFERENCE 1 (bases 1 to 287)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehler, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5435046.
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DXFZ586L2017) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..287
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DXFZ586L2017"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="586 (synonym: hutel)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
ORIGIN
Query Match 5.1%; Score 56; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1034 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 191 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 246
RESULT 14
BI756116
LOCUS
DEFINITION BI756116 290 bp mRNA linear EST 25-SEP-2001
603030013F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200226 5',
mRNA sequence.
ACCESSION
VERSION BI756116
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Homo sapiens
REFERENCE 1 (bases 1 to 290)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@emil.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1A11501 row: j column: 03

High quality sequence stop: 236.

## FEATURES

source

1. 290

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5200226"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_114"

/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 5.1%; Score 56; DB 4; Length 290;

Best Local Similarity 100.0%; Pred. No. 1.3e-15;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1034 ATAGCAAA 1089

Db 138 ATAGCAAA 193

## RESULT 15

AA413331/c

## LOCUS

DEFINITION AGEST00036 Anopheles gambiae adult pSport cDNA Anopheles gambiae

CDNA clone cc38 3', mRNA sequence.

ACCESSION AA413331

VERSION AA413331.1

KEYWORDS GI:2071917

SOURCE EST.

ORGANISM Anopheles gambiae (African malaria mosquito)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;

Anopheles.

1 (bases 1 to 294)

Cornel, A.J., Kumar, V., Mukabayire, O., Salazar Rafferty, C.,

Petrarca, V., Coluzzi, M. and Collins, F.H.

A comprehensive physical map of the malaria vector Anopheles

gambiae

Unpublished (1997)

Other ESTs: AGEST00035

Contact: Salazar Rafferty, C.; and Collins, FH

Vector Genetics Section

Centers for Disease Control and Prevention

MS F-22, 4770 Buford Hwy, Chamblee, GA 30341

Tel: 770 488 7463

Fax: 770 488 7469

Email: czs7@cdc.gov

Mapping location: 34C in 3R chromosome

Seq primer: M13 Universal

High quality sequence stop: 294.

## FEATURES

source

1. 294

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="G3"

/db\_xref="taxon:7165"

/clone="cc38"

/sex="Male and Female"

/clone\_lib="Anopheles gambiae adult pSport cDNA"

/note="Site 1: Sali; Site 2: NotI; See: Salazar, C.E., et

al. Insect Molecular Biology (1994), 3:1-13."

## ORIGIN

Query Match 5.1%; Score 56; DB 1; Length 294;

Best Local Similarity 100.0%; Pred. No. 1.3e-15;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1034 ATAGCAAA 1089

Db 66 ATAGCAAA 11

Search completed: February 6, 2005, 18:50:42

Job time : 4259 secs



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 14:35:16 ; Search time 698 Seconds  
(without alignments)  
9235.814 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcggcagcgcgca.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the chance being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	88.6	1089	4	Aaf80144 Nucleotid
2	168	15.4	168	4	Aaf80149 DNA fragm
3	168	15.4	168	4	Aaf80148 DNA fragm
4	153	14.0	153	4	Aaf80147 DNA fragm
5	150	13.8	150	4	Aaf80150 DNA fragm
6	114	10.5	156	4	Aaf80146 DNA fragm
7	56	5.1	668	5	Aah64916 Human sec
8	55	5.1	376	5	Abv58410 Human pro
9	55	5.1	396	3	Aad02077 cDNA enco
10	55	5.1	410	8	Abx42327 Bovine ES
11	55	5.1	443	4	Aai90926 Human pol
12	55	5.1	491	3	Aaz52570 Human sec
13	55	5.1	917	3	Aac77635 Human can
14	55	5.1	2435	12	Adg22529 Human sof
15	55	5.1	3759	3	Aac58600 Human PRO
16	55	5.1	3819	2	Aaz34292 Human PRO
17	55	5.1	3819	3	Aac78580 Human PRO
18	55	5.1	3819	4	Aas21446 Human CDN
19	55	5.1	3819	6	Abk33602 cDNA enco
20	55	5.1	3819	6	AbL88117 Human PRO

21	55	5.1	3819	6	ABL95606	Human ang
22	55	5.1	3819	8	ACA63860	Novel hum
23	55	5.1	3819	8	ACA03805	cDNA enco
24	55	5.1	3819	8	ACA72024	Human sec
25	55	5.1	3819	8	ABX89343	DNA enco
26	55	5.1	3819	8	ABX92664	cDNA enco
27	55	5.1	3819	8	ACD41997	Human sec
28	55	5.1	3819	8	ACA66405	Human CDN
29	55	5.1	3819	8	ACA68563	Novel hum
30	55	5.1	3819	8	ACA04226	Human CDN
31	55	5.1	3819	9	ADA45924	Novel hum
32	55	5.1	3819	9	ADA76355	Human PRO
33	55	5.1	3819	9	ABT44292	Human PRO
34	55	5.1	3819	9	ADA19005	Human PRO
35	55	5.1	3819	9	ADA61628	Homo sapi
36	55	5.1	3819	9	ADB19413	Novel hum
37	55	5.1	3819	9	ADB27954	cDNA enco
38	55	5.1	3819	9	ADA86433	Novel hum
39	55	5.1	3819	9	ADB15997	Human PRO
40	55	5.1	3819	9	ADA47783	Human PRO
41	55	5.1	3819	9	ADA67578	Human PRO
42	55	5.1	3819	9	ADB30585	cDNA enco
43	55	5.1	3819	9	ADA85881	Novel hum
44	55	5.1	3819	9	ADA97093	Human PRO
45	55	5.1	3819	9	ADA79397	Human PRO

#### ALIGNMENTS

##### RESULT 1

AAF80144

ID AAF80144 standard; DNA; 1089 BP.

AC AAF80144;

DT 11-JUN-2001; (first entry)

DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.

KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;

KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;

OS Triticum monococcum.

FT Key Location/Qualifiers

FT CDS 20..805

FT /\*tag= a

FT /product= "E2F-dimerisation partner (DP) protein"

PN WO200121644-A2.

XX 29-MAR-2001.

XX 25-SEP-2000; 2000WO-EP009325.

XX 24-SEP-1999; 99ES-00002127.

XX 11-NOV-1999; 99ES-00002474.

XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

XX Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/26.

XX P-PSDB; AAB67762.

XX New isolated, enriched, cell free and/or recombinant nucleic acid useful for e.g. altering cell proliferation characteristic such as to alter plant cell, organ or tissue size.

XX Claim 13; Fig 1; 77pp; English.

CC The present sequence encodes a E2F-dimerisation partner (DP) protein. The  
 CC proteins acts as a plant E2F transcription factor. E2F and DP are two  
 CC proteins that hetero-dimerise to form an active transcription factor that  
 CC regulates G1 to S phase of the cell cycle, and later, the expression of  
 CC genes required for S-phase progression. E2F and retinoblastoma protein  
 CC also interact as a hetero-dimer in cells to suppress certain genes. This  
 CC repression involves binding of the retinoblastoma protein to the E2F-DP  
 CC dimer that is in turn bound to sites on DNA through the E2F DNA binding  
 CC domain. DP proteins can be modulated to alter plant cell, organ or tissue  
 CC shape, and particularly to alter cell proliferation characteristic such  
 CC as to alter plant cell, organ or tissue size  
 XX  
 SQ Sequence 1089 BP; 367 A; 215 C; 261 G; 246 T; 0 U; 0 Other;

Query Match 88.6%; Score 965; DB 4; Length 1089;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GCGGTAATCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184  
 DB 125 GCGGTAATCGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGGAG 184

QY 185 AAGGCTGCGGACCGAGGATCAACGGTTGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244  
 DB 185 AAGGCTGCGGACCGAGGATCAACGGTTGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244

QY 245 GAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCAGCAAGAAATTTATCA 304  
 DB 245 GAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCAGCAAGAAATTTATCA 304

QY 305 GAGCTGAAGTCCATGCGACATATTTGTCAGGGTTTGTGAGAGAAATATAGCGCGAGA 364  
 DB 305 GAGCTGAAGTCCATGCGACATATTTGTCAGGGTTTGTGAGAGAAATATATAGCGCGAGA 364

QY 365 GTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTGTTATTCGAAAGAAAAAGGAG 424  
 DB 365 GTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTGTTATTCGAAAGAAAAAGGAG 424

QY 425 ATACGTTGAGTGGGCTTTCAAAATTCAGATATGAAAGAAATTAAGAGCTTGAGGAAGTT 484  
 DB 425 ATACGTTGAGTGGGCTTTCAAAATTCAGATATGAAAGAAATTAAGAGCTTGAGGAAGTT 484

QY 485 CGTAAAGAACTCGTCAACAGATTAGCAACAAGAGCACTCTCCAGGAAATCGAAAAA 544  
 DB 485 CGTAAAGAACTCGTCAACAGATTAGCAACAAGAGCACTCTCCAGGAAATCGAAAAA 544

QY 545 CAGTTTATGATGATCTCCAAACATCAAGTTACGTAACCAACACTTGGAAAGCTCAGCAGAG 604  
 DB 545 CAGTTTATGATGATCTCCAAACATCAAGTTACGTAACCAACACTTGGAAAGCTCAGCAGAG 604

QY 605 AATGTTAATGGCATCCGCTTCATTCGTTATTTGGTCAAGACATCTAGGAAAGCAAGGGTG 664  
 DB 605 AATGTTAATGGCATCCGCTTCATTCGTTATTTGGTCAAGACATCTAGGAAAGCAAGGGTG 664

QY 665 GAAATTTGAGATTTTCAGATGCTCGAAGTTTGGCCATTCAGTTCAATGTTGACCAATTC 724  
 DB 665 GAAATTTGAGATTTTCAGATGCTCGAAGTTTGGCCATTCAGTTCAATGTTGACCAATTC 724

QY 725 ACATTTGATGATGATCTCTCAATCTTTAGGGGGTAAAGCGTAAACAGCATAGGAAGAGCT 784  
 DB 725 ACATTTGATGATGATCTCTCAATCTTTAGGGGGTAAAGCGTAAACAGCATAGGAAGAGCT 784

QY 785 GGGCGGCGCCATCTTCATAGAGATCAAGAAATATTACAAATTAAGAAAGTTAGAA 844  
 DB 785 GGGCGGCGCCATCTTCATAGAGATCAAGAAATATTACAAATTAAGAAAGTTAGAA 844

QY 845 CTGGCAGCGCGATCTTTTGGCACAGCTATGATAGCTATATATCTCATGAAACTTG 904  
 DB 845 CTGGCAGCGCGATCTTTTGGCACAGCTATGATAGCTATATATCTCATGAAACTTG 904

QY 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAAGATTTTAACTGCAAAATTTGCTC 964  
 DB 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAAGATTTTAACTGCAAAATTTGCTC 964

QY 965 CTTTTTGTGCTAGCAGGTTATTAGCTCTCAGATAGATGATTCATATATGCTGCTATG 1024  
 DB 965 CTTTTTGTGCTAGCAGGTTATTAGCTCTCAGATAGATGATTCATATATGCTGCTATG 1024

QY 1025 AAAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084  
 DB 1025 AAAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084

QY 1085 AAAAA 1089  
 DB 1085 AAAAA 1089

RESULT 2  
 AAF80149  
 ID AAF80149 standard; DNA; 168 BP.  
 XX  
 AC AAF80149;  
 XX  
 DT 11-JUN-2001 (first entry)  
 XX  
 DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.  
 XX  
 KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
 KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
 KW ss.  
 XX  
 OS Triticum monococcum.  
 XX  
 PN WO200121644-A2.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PF 25-SEP-2000; 2000WO-EP009325.  
 XX  
 PR 24-SEP-1999; 99ES-00002127.  
 PR 11-NOV-1999; 99ES-00002474.  
 XX  
 PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX  
 PI Gutierrez-Armenta C, Ramirez-Parra E;  
 XX  
 DR WPI; 2001-257972/26.  
 DR P-PSDB; AAB67767.  
 XX  
 PT New isolated, enriched, cell free and/or recombinant nucleic acid useful  
 PT for e.g. altering cell proliferation characteristic such as to alter  
 PT plant cell, organ or tissue size.  
 XX  
 PS Disclosure; Page 74-75; 77pp; English.  
 XX  
 CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation  
 CC partner (DP) protein. The protein acts as a plant E2F transcription  
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an  
 CC active transcription factor that regulates G1 to S phase of the cell  
 CC cycle, and later, the expression of genes required for S-phase  
 CC progression. E2F and retinoblastoma protein also interact as a hetero-  
 CC dimer in cells to suppress certain genes. This repression involves  
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn  
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and particularly  
 CC to alter cell proliferation characteristic such as to alter plant cell,  
 CC organ or tissue size  
 XX  
 SQ Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;

Query Match 15.4%; Score 168; DB 4; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-49;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 CGTAAAGAACTCGTCAACAGATTAGCAACAAGAGCACTCTCCAGGAATCGAAAAA 544  
 |||||

```
Db 1 CGTAAAGAACTCGTCAACAGATTAGGAAACAAGAGGCACTCTCTCAGGAAATCGAAAAA 60
QY 545 CAGTTTGATGATCTCCAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 604
Db 61 CAGTTTGATGATCTCCAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 120
QY 605 AATGTTAATGGCATCGCCCTTCATTGCGTATTGGTCAAGACATCTAGG 652
Db 121 AATGTTAATGGCATCGCCCTTCATTGCGTATTGGTCAAGACATCTAGG 168

RESULT 3
AAF80148
ID AAF80148 standard; DNA; 168 BP.
XX AC AAF80148;
XX DT 11-JUN-2001 (first entry)
XX DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
XX KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
XX KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
XX OS Triticum monococcum.
XX PN WO200121644-A2.
XX PD 29-MAR-2001.
XX PF 25-SEP-2000; 2000WO-EP009325.
XX PR 24-SEP-1999; 99ES-00002127.
XX PR 11-NOV-1999; 99ES-00002474.
XX PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX PI Gutierrez-Armenta C, Ramirez-Parra E;
XX DR WPI; 2001-257972/26.
XX DR P-PSDB; AAB67766.
XX PT New isolated, enriched, cell free and/or recombinant nucleic acid useful
XX PT for e.g. altering cell proliferation characteristic such as to alter
XX PT plant cell, organ or tissue size.
XX PS Disclosure; Page 73-74; 77pp; English.
XX CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
XX CC partner (DP) protein. The protein acts as a plant E2F transcription
XX CC factor. E2F and DP are two proteins that hetero-dimerise to form an
XX CC active transcription factor that regulates G1 to S phase of the cell
XX CC cycle, and later, the expression of genes required for S-phase
XX CC progression. E2F and retinoblastoma protein also interact as a hetero-
XX CC dimer in cells to suppress certain genes. This repression involves
XX CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
XX CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
XX CC be modulated to alter plant cell, organ or tissue shape, and particularly
XX CC to alter cell proliferation characteristic such as to alter plant cell,
XX CC organ or tissue size
XX SQ Sequence 168 BP; 59 A; 19 C; 44 G; 46 T; 0 U; 0 Other;

Query Match 15.4%; Score 168; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.8e-49;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGTTTGATGAGAGAATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGTTTGATGAGAGAATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGCACTTCGTGTTATTGCAAAAGAAAAAAGGAGATACGGTGGATG 436
```

```
Db 61 TTCAACGTTCTCATTCGCACTTCGTGTTATTGCAAAAGAAAAAAGGAGATACGGTGGATG 120
QY 437 GGCCTTTTCAAAATACAGATATGAAAAAATAAAGAGCTTTGAGGAAGTT 484
Db 121 GGCCTTTTCAAAATACAGATATGAAAAAATAAAGAGCTTTGAGGAAGTT 168

RESULT 4
AAF80147
ID AAF80147 standard; DNA; 153 BP.
XX AC AAF80147;
XX DT 11-JUN-2001 (first entry)
XX DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
XX KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
XX KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
XX OS Triticum monococcum.
XX PN WO200121644-A2.
XX PD 29-MAR-2001.
XX PF 25-SEP-2000; 2000WO-EP009325.
XX PR 24-SEP-1999; 99ES-00002127.
XX PR 11-NOV-1999; 99ES-00002474.
XX PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX PI Gutierrez-Armenta C, Ramirez-Parra E;
XX DR WPI; 2001-257972/26.
XX DR P-PSDB; AAB67765.
XX PT New isolated, enriched, cell free and/or recombinant nucleic acid useful
XX PT for e.g. altering cell proliferation characteristic such as to alter
XX PT plant cell, organ or tissue size.
XX PS Disclosure; Page 72-73; 77pp; English.
XX CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
XX CC partner (DP) protein. The protein acts as a plant E2F transcription
XX CC factor. E2F and DP are two proteins that hetero-dimerise to form an
XX CC active transcription factor that regulates G1 to S phase of the cell
XX CC cycle, and later, the expression of genes required for S-phase
XX CC progression. E2F and retinoblastoma protein also interact as a hetero-
XX CC dimer in cells to suppress certain genes. This repression involves
XX CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
XX CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
XX CC be modulated to alter plant cell, organ or tissue shape, and particularly
XX CC to alter cell proliferation characteristic such as to alter plant cell,
XX CC organ or tissue size
XX SQ Sequence 153 BP; 55 A; 25 C; 47 G; 26 T; 0 U; 0 Other;

Query Match 14.0%; Score 153; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 3.4e-44;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 GATAAGATAGGAAGAAGGAGGCTCGGCAACGAGGATCACCGTTGGGGGCTCCGC 223
Db 1 GATAAGATAGGAAGAAGGAGGCTCGGCAACGAGGATCACCGTTGGGGGCTCCGC 60
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATCATATGAG 283
Db 61 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATCATATGAG 120
```

QY 773 ATAGGAAGAGCTGGCCGCGCCACCCCTTCAC 802  
|||  
pb 121 ATAGGAAGAGCTGGCCGCGCCACCCCTTCAC 150

```
XX 11-SEP-2001 (first entry)
XX Human secreted protein cDNA, SEQ ID NO: 192.
XX
XX Human secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX GENSET; ss.
XX
XX Homo sapiens.
XX
XX WO200142451-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-1B001938.
XX
XX 08-DEC-1999; 99US-0169629P.
XX
XX 06-MAR-2000; 2000US-0187470P.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
XX WPI; 2001-367870/38.
XX
XX P-PSDB; AAC89313.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
XX proteins, useful in gene therapy and vaccination against a variety of
XX diseases, and for diagnosis of those diseases.
XX
XX Claim 7; Page 753; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased GENSET
XX gene expression by rectifying mutations or deletions in a patient's
XX genome that affect the activity of GENSET or by supplementing the
XX patient's own production of GENSET polypeptides. Conversely, antisense
XX nucleic acid molecules may be administered to down regulate GENSET
XX expression by binding with the cells' own genes and preventing their
XX expression. The sense and antisense nucleic acids may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acid sequences in samples, and hence to determine which
XX patients may be in need of restorative therapy. The GENSET polypeptides
XX may be used as antigens in the production of antibodies and in assays to
XX identify modulators (agonists and antagonists) of GENSET polypeptide
XX expression and activity. The present sequence is a GENSET nucleic acid of
XX the invention
XX
XX Sequence 668 BP; 188 A; 163 C; 151 G; 166 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 56; DB 5; Length 668;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-10;
XX Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1034 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX DB 594 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 649
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX
XX RESULT 8
XX ABV58410
XX ID ABV58410 standard; cDNA; 376 BP.
XX
XX AC ABV58410;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 58401.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
```

```
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JB;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 11212; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 376 BP; 164 A; 68 C; 61 G; 83 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 5; Length 376;
XX Best Local Similarity 100.0%; Pred. No. 6.8e-10;
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX DB 215 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 269
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX
XX RESULT 9
XX AAD02077
XX ID AAD02077 standard; cDNA; 396 BP.
XX
XX AC AAD02077;
XX
XX 26-MAR-2001 (first entry)
XX
XX cDNA encoding rat pituitary hormone, pituitrone.
XX
XX Rat; pituitary; pituitrone; therapy; immune disorder; anaemia;
XX Digorge syndrome; haematopoietic cell; ataxia telangiectasia;
XX Wiskott-Aldrich disorder; blood coagulation disorder; Addison's disease;
XX autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
XX SLE; hyperproliferative disorder; gene therapy; neoplasm;
XX infectious disease; immunomodulatory; cytostatic; antimicrobial; ss.
XX
XX Rattus sp.
XX
XX Key Location/Qualifiers
XX FH
```

```

FT CDS 1. .210
FT /*tag= a
FT /products= "Rat pituitary hormone, pituitrone"
FT /note= "Does not include start codon"
FT /partial
XX
XX WO200066778-A1.
XX
XX 09-NOV-2000.
XX
XX 27-APR-2000; 2000WO-US011211.
XX
XX 30-APR-1999; 99US-0131966P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J;
XX
XX WPI; 2000-687547/67.
XX
XX P-PSDB; AAY71961.
XX
XX Nucleic acids encoding a novel pituitary hormone (pituitrone), useful for
XX diagnosing, preventing and treating e.g. immune disorders,
XX hyperproliferative disorders and blood coagulation disorders.
XX
XX Claim 1; Page 269; 277pp; English.
XX
XX The present sequence is a cDNA encoding rat pituitary hormone,
XX pituitrone. Pituitrone is highly expressed in pituitary gland and is also
XX expressed in brain tissues, spinal cord and kidney. It may be used as
XX antigens in the production of antibodies against pituitrone and in assays
XX to identify modulators. Pituitrone cDNAs are also useful in gene therapy.
XX Pituitrone may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate pituitrone expression. It may be
XX useful in treating disorders related to reproductive and renal system,
XX immune disorders, disorders of haematopoietic cells (e.g., anaemia,
XX George syndrome, ataxia telangiectasia and Wiskott-Aldrich disorder),
XX blood coagulation disorders, autoimmune disorders (e.g. Addison's
XX disease, multiple sclerosis and systemic lupus erythematosus (SLE)),
XX hyperproliferative disorders (e.g. neoplasms of bone, liver and pancreas)
XX and infectious diseases
XX
XX Sequence 396 BP; 110 A; 135 C; 91 G; 60 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 3; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-10;
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX DB 342 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 396
XX
XX RESULT 10
XX ID ABX42327/c
XX ABX42327 standard; cDNA; 410 BP.
XX
XX AC ABX42327;
XX
XX XX 20-FEB-2003 (first entry)
XX
XX DE Bovine EST associated with lactation/muscle/fat deposition #7492.
XX
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
XX OS Bos Taurus.
XX
XX PN US2002137139-A1.
XX
XX XX 26-SEP-2002.
XX
XX

```

---

```

PF 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
XX 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 7492; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMPD), derived from
XX cattle, and the LMPD nucleic acid can specifically hybridise to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMPD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
XX Sequence 410 BP; 151 A; 73 C; 60 G; 126 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 8; Length 410;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-10;
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX DB 80 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 26
XX
XX RESULT 11
XX ID AAI90926
XX AAI90926 standard; cDNA; 443 BP.
XX
XX AC AAI90926;
XX
XX XX 06-NOV-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 10986.
XX
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX

```



```

OS Homo sapiens.
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56;
XX
XX P-PSDB; AAO10995.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 10986; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX
XX the encoded proteins (AA00010-AA013910) that exhibit activity elating to
XX
XX cytokine, cell proliferation or cell differentiation or which may induce
XX
XX production of other cytokines in other cell populations. The
XX
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX
XX activity, tissue growth factor activity, immunomodulatory activity and
XX
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX
XX inflammation. Note: The sequence data for this patent did not form part
XX
XX of the printed specification, but was obtained in electronic format
XX
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 443 BP; 183 A; 68 C; 96 G; 96 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 4; Length 443;
XX
XX Best Local Similarity 100.0%; Pred. NO. 6.5e-10;
XX
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX
XX Db 137 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 191
XX
XX RESULT 12
XX
XX AAZ52570
XX
XX ID AAZ52570 standard; cDNA; 491 BP.
XX
XX XX
XX
XX AC AAZ52570;
XX
XX XX
XX
XX 29-FEB-2000 (first entry)
XX
XX
XX Human secreted protein clone yd261_1 nucleotide sequence SEQ ID NO:191.
XX
XX
XX Human; secreted protein; immunostimulatory; haemostatic; cytokine;
XX
XX proliferative; differentiative; chemotactic; chemokinetic; vaccine;
XX
XX thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
XX
XX gene therapy; ss.
XX
XX
XX Homo sapiens.
XX
XX XX
XX
XX WO9958642-A2.
XX
XX
XX 18-NOV-1999.
XX
XX
XX 14-MAY-1999; 99WO-US010843.
XX
XX
XX 14-MAY-1998; 98US-0085472P.
XX
XX
XX 17-AUG-1998; 98US-0096824P.
XX

```

```

PR 11-SEP-1998; 98US-0099843P.
PR
PR 11-SEP-1998; 98US-0099950P.
PR
PR 15-SEP-1998; 98US-0100424P.
PR
PR 29-SEP-1998; 98US-0102329P.
PR
PR 09-OCT-1998; 98US-0103615P.
PR
PR 11-DEC-1998; 98US-0111779P.
PR
PR 14-DEC-1998; 98US-0112159P.
PR
PR 31-DEC-1998; 98US-0114415P.
PR
PR 10-FEB-1999; 99US-00248059.
PR
PR 06-APR-1999; 99US-00287150.
PR
PR 13-MAY-1999; 99US-00311021.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX PA
XX
XX PI Wong GG, Clark HF, Fechtel K, Agostino MJ;
XX
XX XX
XX
XX DR WPI; 2000-053095/04.
XX
XX DR P-PSDB; AAY73485.
XX
XX
XX Novel polynucleotides and proteins having biological activities which
XX
XX make them suitable for treating, preventing or ameliorating medical
XX
XX conditions in humans or animals.
XX
XX
XX Claim 200; Page 702; 730pp; English.
XX
XX The present invention describes human secreted proteins encoded by
XX
XX polynucleotides obtained from adult testes, foetal brain, adult brain,
XX
XX brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
XX
XX cDNA libraries. The polynucleotides and proteins are predicted to have
XX
XX biological activities which would make them suitable for treating,
XX
XX preventing or ameliorating medical conditions in humans and animals.
XX
XX Suggested activities include nutritional activity, cytokine and cell
XX
XX proliferation/differentiation activity, immune stimulating (e.g. as
XX
XX vaccines) or suppressing activity, haematopoiesis regulating activity,
XX
XX tissue growth activity, activin/inhibin activity, chemotactic/
XX
XX chemokinetic activity, haemostatic and thrombolytic activity, receptor/
XX
XX ligand activity, anti-inflammatory activity, cadherin/tumour invasion
XX
XX suppressor activity, and tumour inhibition activity. The polynucleotides
XX
XX are also stated to be useful for gene therapy. Therapeutic compositions
XX
XX are also presently valuable for veterinary applications. AAZ52475 to
XX
XX AAZ52581 encode human secreted proteins, and AAY73390 to AAY73500
XX
XX represent human secreted proteins, given in the present invention
XX
XX
XX Sequence 491 BP; 133 A; 148 C; 114 G; 96 T; 0 U; 0 Other;
XX
XX
XX Query Match 5.1%; Score 55; DB 3; Length 491;
XX
XX Best Local Similarity 100.0%; Pred. NO. 6.4e-10;
XX
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX
XX Db 430 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 484
XX
XX
XX RESULT 13
XX
XX AAC77635
XX
XX ID AAC77635 standard; cDNA; 917 BP.
XX
XX XX
XX
XX AC AAC77635;
XX
XX XX
XX
XX 08-FEB-2001 (first entry)
XX
XX
XX DE Human cancer associated gene sequence SEQ ID NO:29.
XX
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX
XX diagnosis; cytostatic; proliferative; cancer antigen; immunomodulator;
XX
XX antidiabetic; antiaesthatic; antirheumatic; antithrombotic; antiviral;
XX
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
XX
XX dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
XX
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX
XX allergic reaction; graft versus host disease; organ rejection;
XX
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX

```

KW neurological disease; drug screening; ss.  
 XX Homo sapiens.  
 OS WO200055350-A1.  
 PN 21-SEP-2000.  
 PD 08-MAR-2000; 2000WO-US005882.  
 XX 12-MAR-1999; 99US-0124270P.  
 XX (HUMA-) HUMAN GENOME SCI INC..  
 PA Rosen CA, Ruben SM;  
 PI WPI; 2000-587533/55.  
 DR P-PSDB; AAB43426.  
 XX Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer.  
 PT Claim 1; Page 632; 2352pp; English.  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given in  
 CC AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnary; immunomodulator;  
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC nootropic; vasotropic; antipeptidic; antidiabetic; antihypertensive; The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention  
 XX  
 SQ Sequence 917 BP; 247 A; 269 C; 257 G; 142 T; 0 U; 2 Other;  
 Query Match 5.1%; Score 55; DB 3; Length 917;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1035 TAGCAAA 1089  
 DB 841 TAGCAAA 895  
 RESULT 14  
 ADQ22529  
 ID ADQ22529 standard; DNA; 2435 BP.  
 XX  
 AC ADQ22529;  
 XX  
 XX 26-AUG-2004 (first entry)  
 DT Human soft tissue sarcoma-upregulated DNA - SEQ ID 5349.  
 DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
 XX db.  
 KW Homo sapiens.  
 OS  
 XX  
 XX WO2004048938-A2.  
 PN

10-JUN-2004.  
 26-NOV-2003; 2003WO-US038193.  
 26-NOV-2002; 2002US-0429739P.  
 (PROT-) PROTEIN DESIGN LABS INC.  
 Aziz N, Ginsburg WM, Zlotnik A;  
 WPI; 2004-441208/41.  
 Early detection of soft tissue sarcoma comprises determining expression  
 of a gene in a first soft tissue sample and a normal soft tissue sample  
 and comparing the gene expression, also useful in treating soft tissue  
 sarcoma.  
 Example 2; SEQ ID NO 5349; 210pp; English.  
 The invention relates to a novel method for detecting soft tissue sarcoma  
 which comprises obtaining a first soft tissue sample from an individual  
 and a normal soft tissue sample from the same or different individual,  
 determining the expression of a gene in both samples and comparing the  
 expression of the gene in both soft tissue samples, where a higher level  
 of protein expression in the first soft tissue sample indicates the  
 presence of soft tissue sarcoma. The method of the invention has  
 cytostatic applications and may be useful for detecting soft tissue  
 sarcoma, possibly via gene therapy or vaccine production. The nucleic  
 acid sequences may be useful in diagnostic and screening applications.  
 The current sequence is that of a human soft tissue sarcoma-upregulated  
 DNA of the invention. The current sequence is not shown within the  
 specification per se but was submitted in CD format by the inventor.  
 Sequence 2435 BP; 637 A; 571 C; 538 G; 688 T; 0 U; 1 Other;  
 Query Match 5.1%; Score 55; DB 12; Length 2435;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1035 TAGCAAA 1089  
 DB 2358 TAGCAAA 2412  
 RESULT 15  
 AAC58600  
 ID AAC58600 standard; cDNA; 3759 BP.  
 XX  
 AC AAC58600;  
 XX  
 DT 29-JAN-2001 (first entry)  
 XX Human PRO1083 protein UNQ540 encoding cDNA SEQ ID NO:116.  
 DE  
 XX  
 XX Human; immune related disease; diagnosis; antiinflammatory; cardiant;  
 KW dermatological; antiarthritic; antirheumatic; immunosuppressive;  
 KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;  
 KW antinaemic; hepatotropic; viricide; antipsoriatic; antiallergic;  
 KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;  
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;  
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;  
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;  
 KW autoimmune thrombocytopaenia; immune-mediated renal disease;  
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;  
 KW inflammatory bowel disease; gluten-sensitive enteropathy;  
 KW autoimmune disease; immune-mediated skin disease; allergic disease;  
 KW immunological disease; transplantation associated disease;  
 KW graft rejection; graft-versus-host-disease; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200053758-A2.  
 PN

XX 14-SEP-2000.  
 XX 02-MAR-2000; 2000WO-US005841.  
 XX 08-MAR-1999; 99WO-US005028.  
 XX 10-MAR-1999; 99US-0123618P.  
 XX 12-MAR-1999; 99US-0123957P.  
 XX 23-MAR-1999; 99US-0125775P.  
 XX 12-APR-1999; 99US-0128849P.  
 XX 20-APR-1999; 99WO-US008615.  
 XX 28-APR-1999; 99US-0131445P.  
 XX 04-MAY-1999; 99US-0132371P.  
 XX 14-MAY-1999; 99US-0134287P.  
 XX 02-JUN-1999; 99WO-US012252.  
 XX 23-JUN-1999; 99US-0141037P.  
 XX 26-JUL-1999; 99US-0144758P.  
 XX 28-JUL-1999; 99US-0145698P.  
 XX 01-SEP-1999; 99WO-US020111.  
 XX 08-SEP-1999; 99WO-US020594.  
 XX 13-SEP-1999; 99WO-US020944.  
 XX 15-SEP-1999; 99WO-US021090.  
 XX 05-OCT-1999; 99WO-US021547.  
 XX 29-OCT-1999; 99WO-US023089.  
 XX 29-NOV-1999; 99US-0162506P.  
 XX 30-NOV-1999; 99WO-US028214.  
 XX 30-NOV-1999; 99WO-US028313.  
 XX 01-DEC-1999; 99WO-US028409.  
 XX 01-DEC-1999; 99WO-US028301.  
 XX 02-DEC-1999; 99WO-US028634.  
 XX 02-DEC-1999; 99WO-US028551.  
 XX 02-DEC-1999; 99WO-US028564.  
 XX 16-DEC-1999; 99WO-US028565.  
 XX 20-DEC-1999; 99WO-US030095.  
 XX 30-DEC-1999; 99WO-US030999.  
 XX 05-JAN-2000; 99WO-US031274.  
 XX 06-JAN-2000; 2000WO-US000219.  
 XX 06-JAN-2000; 2000WO-US000277.  
 XX 11-FEB-2000; 2000WO-US000376.  
 XX 18-FEB-2000; 2000WO-US003565.  
 XX 18-FEB-2000; 2000WO-US004341.  
 XX 22-FEB-2000; 2000WO-US004342.  
 XX 22-FEB-2000; 2000WO-US004414.  
 (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;  
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;  
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;  
 DR WPI: 2000-572271/53.  
 DR P-PSDB; AAB33435.

XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of  
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid  
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus.

XX Claim 23; Fig 43; 309pp; English.

XX The present invention describes sixty four human PRO proteins which can  
 CC be used in the treatment of immune related diseases. The human PRO  
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for  
 CC treating and diagnosing immune related disorders. The disorders are  
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,  
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis,  
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's  
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic  
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,  
 CC immune-mediated renal disease, demyelinating diseases of the central and  
 CC peripheral nervous systems, hepatobiliary diseases, inflammatory bowel  
 CC disease, gluten-sensitive enteropathy and Whipple's disease, autoimmune  
 CC or immune-mediated skin diseases, allergic diseases, immunological

CC diseases of the lung, and transplantation associated diseases including  
 CC graft rejection and graft-versus-host-disease. AAC58397 to AAC58578  
 CC represent PCR primers and hybridisation probes used in the isolation of  
 CC human PRO sequences. AAC58579 to AAC58642 and AAB33414 to AAB33477  
 CC represent human PRO polynucleotide and protein sequences given in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 3759 BP; 774 A; 1205 C; 970 G; 810 T; 0 U; 0 Other;  
 Query Match 5.1%; Score 55; DB 3; Length 3759;  
 Best Local Similarity 100.0%; Pred. No. 4e-10;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1035 TAGCAAA 1089  
 Db 3675 TAGCAAA 3729

Search completed: February 6, 2005, 17:35:48  
 Job time : 704 secs

***This Page Blank (usp10)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 17:39:42 ; Search time 690 Seconds  
(without alignments)

9084.062 Million cell updates/sec

Title: us-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgacgcgcgcgca.....aaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	5.4	1199	18	US-10-425-115-36465
2	56	5.1	668	9	US-09-731-872-192
3	56	5.1	668	10	US-09-876-997-192
4	56	5.1	2121	18	US-10-425-115-184240
5	55	5.1	222	18	US-10-425-115-150417
6	55	5.1	369	18	US-10-425-115-52629
7	55	5.1	376	18	US-10-357-930-58429
8	55	5.1	410	9	US-09-360-352-7492
9	55	5.1	538	18	US-10-425-115-91860
10	55	5.1	787	17	US-10-424-599-122506
11	55	5.1	838	18	US-10-425-115-163680

12	55	5.1	917	9	US-09-925-301-29	Sequence 29, Appl
13	55	5.1	1027	9	US-09-803-589-13	Sequence 13, Appl
14	55	5.1	1027	18	US-10-718-332-13	Sequence 13, Appl
15	55	5.1	2435	18	US-10-723-860-5349	Sequence 5349, Ap
16	55	5.1	3819	9	US-09-978-295A-482	Sequence 482, App
17	55	5.1	3819	9	US-09-978-697-482	Sequence 482, App
18	55	5.1	3819	9	US-09-978-192A-482	Sequence 482, App
19	55	5.1	3819	9	US-09-999-832A-482	Sequence 482, App
20	55	5.1	3819	10	US-09-978-189-482	Sequence 482, App
21	55	5.1	3819	10	US-09-978-608A-482	Sequence 482, App
22	55	5.1	3819	10	US-09-978-585A-482	Sequence 482, App
23	55	5.1	3819	10	US-09-978-191A-482	Sequence 482, App
24	55	5.1	3819	10	US-09-978-403A-482	Sequence 482, App
25	55	5.1	3819	10	US-09-978-564A-482	Sequence 482, App
26	55	5.1	3819	10	US-09-999-833A-482	Sequence 482, App
27	55	5.1	3819	10	US-09-981-915A-482	Sequence 482, App
28	55	5.1	3819	10	US-09-978-824-482	Sequence 482, App
29	55	5.1	3819	10	US-09-918-585A-482	Sequence 482, App
30	55	5.1	3819	10	US-09-999-834A-482	Sequence 482, App
31	55	5.1	3819	10	US-09-978-423A-482	Sequence 482, App
32	55	5.1	3819	10	US-09-978-193A-482	Sequence 482, App
33	55	5.1	3819	10	US-09-999-830A-482	Sequence 482, App
34	55	5.1	3819	10	US-09-978-757A-482	Sequence 482, App
35	55	5.1	3819	10	US-09-978-187B-482	Sequence 482, App
36	55	5.1	3819	10	US-09-978-643A-482	Sequence 482, App
37	55	5.1	3819	10	US-09-978-375A-482	Sequence 482, App
38	55	5.1	3819	10	US-09-978-298A-482	Sequence 482, App
39	55	5.1	3819	10	US-09-978-188A-482	Sequence 482, App
40	55	5.1	3819	10	US-09-978-681A-482	Sequence 482, App
41	55	5.1	3819	10	US-09-978-194A-482	Sequence 482, App
42	55	5.1	3819	10	US-09-999-829A-482	Sequence 482, App
43	55	5.1	3819	10	US-09-978-299A-482	Sequence 482, App
44	55	5.1	3819	10	US-09-978-544A-482	Sequence 482, App
45	55	5.1	3819	10	US-09-978-665A-482	Sequence 482, App

#### ALIGNMENTS

RESULT 1  
US-10-425-115-36465  
; Sequence 36465, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 36465  
; LENGTH: 1199  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_133259C.1  
US-10-425-115-36465

Query Match	5.4%	Score 59	DB 18	Length 1199
Best Local Similarity	100.0%	Pred. No. 4.9e-19	Indels 0	Gaps 0
Matches 59	Conservative 0	Mismatches 0		
QY	1031	TTGATAGCAAA	1089	
DB	738	TTGATAGCAAA	796	
RESULT 2				
US-09-731-872-192				

## RESULT 4

```

RESULT 6
US-10-425-115-52629
; Sequence 52629, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 52629  
; LENGTH: 369  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_147996C.1  
US-10-425-115-52629

Query Match 5.1%; Score 55; DB 18; Length 376;  
Best Local Similarity 100.0%; Pred. No. 5.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089  
Db 129 TAGCAAA 183

## RESULT 7

US-10-357-930-58429  
; Sequence 58429, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007ECN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 58429  
; LENGTH: 376  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-357-930-58429

Query Match 5.1%; Score 55; DB 18; Length 376;  
Best Local Similarity 100.0%; Pred. No. 5.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089  
Db 215 TAGCAAA 269

## RESULT 8

US-09-960-352-7492/C  
; Sequence 7492, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 7492  
; LENGTH: 410  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 32-LIB34-017-Q1-E1-H7  
US-09-960-352-7492

Query Match 5.1%; Score 55; DB 9; Length 410;  
Best Local Similarity 100.0%; Pred. No. 5.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089  
Db 80 TAGCAAA 26

## RESULT 9

US-10-425-115-91860/C  
; Sequence 91860, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 91860  
; LENGTH: 538  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_183771C.1  
US-10-425-115-91860

Query Match 5.1%; Score 55; DB 18; Length 538;  
Best Local Similarity 100.0%; Pred. No. 5.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089  
Db 96 TAGCAAA 42

## RESULT 10

US-10-424-599-122506  
; Sequence 122506, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 122506

```
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81629C.1
US-10-424-599-122506

Query Match          5.1%; Score 55; DB 17; Length 787;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 672 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 726

RESULT 11
US-10-425-115-163680
; Sequence 163680, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 163680
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(838)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80852C.1
US-10-425-115-163680

Query Match          5.1%; Score 55; DB 18; Length 838;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 670 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 12
US-09-925-301-29
; Sequence 29, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-29
```

```
Query Match          5.1%; Score 55; DB 9; Length 917;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 841 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 895

RESULT 13
US-09-803-589-13
; Sequence 13, Application US/09803589
; Patent No. US20020112251A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-325001
; CURRENT APPLICATION NUMBER: US/09/803,589
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(630)
US-09-803-589-13

Query Match          5.1%; Score 55; DB 9; Length 1027;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 958 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

RESULT 14
US-10-718-332-13
; Sequence 13, Application US/10718332
; Publication No. US20040253605A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-325001
; CURRENT APPLICATION NUMBER: US/10/718,332
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/803,589
```



;; PRIOR FILING DATE: 2001-03-09  
;; PRIOR APPLICATION NUMBER: US 09/128,709  
;; PRIOR FILING DATE: 1998-08-04  
;; PRIOR APPLICATION NUMBER: US 60/054,645  
;; PRIOR FILING DATE: 1997-08-04  
;; PRIOR APPLICATION NUMBER: US 09/130,491  
;; PRIOR FILING DATE: 1998-08-06  
;; PRIOR APPLICATION NUMBER: US 60/054,966  
;; PRIOR FILING DATE: 1997-08-06  
;; PRIOR APPLICATION NUMBER: US 60/058,108  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: US 09/388,280  
;; PRIOR FILING DATE: 1999-09-01  
;; PRIOR APPLICATION NUMBER: US 09/388,279  
;; PRIOR FILING DATE: 1999-09-01  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 1027  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (106)...(630)  
US-10-718-332-13

Query Match 5.1%; Score 55; DB 18; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 5.2e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1035 TAGCAAA 1089  
DB 958 TAGCAAA 1012

RESULT 15  
US-10-723-860-5349  
; Sequence 5349, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; FILE REFERENCE: 05892.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5349  
; LENGTH: 2435  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2063)..(2063)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-723-860-5349

Query Match 5.1%; Score 55; DB 18; Length 2435;  
Best Local Similarity 100.0%; Pred. No. 5.2e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1035 TAGCAAA 1089  
DB 2358 TAGCAAA 2412

Search completed: February 6, 2005, 21:54:29  
Job time : 692 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:47:21 ; Search time 5146 Seconds  
(without alignments)  
10254.126 Million cell updates/sec

Title: US-10-088-830-1  
Perfect score: 1089  
Sequence: 1 gaattcgacgagccgcaa.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.2	99.6	1089	6	AX100704
2	1065.4	97.8	1083	8	TSP271917
3	515	47.3	1174	8	AK111611
4	442.2	40.6	885	8	AY224529
5	168	15.4	188	6	AX100712
6	168	15.4	188	6	AX100714
7	153.6	14.1	879	8	BT005286
8	153.6	14.1	960	8	ATH319027
9	153.6	14.1	1114	6	AX449297
10	153.6	14.1	1121	8	AK117135
11	153.6	14.1	1189	6	AX449329
12	153.6	14.1	1274	6	AX712207
13	153.6	14.1	1274	6	AX734251
14	153.6	14.1	1274	8	ATH294531
15	153	14.0	153	6	AX100710
16	151.2	13.9	156	6	AX100708
17	150.8	13.8	1158	8	ATH294532
18	150.8	13.8	1618	8	AY086018
19	150.6	13.8	1640	8	AF181998

20	150.6	13.8	1640	8	AY307373
21	150	13.8	150	6	AX100716
22	136.2	12.5	1140	8	AY224589
23	132.2	12.1	1041	8	AY224551
24	120.6	11.1	19459	8	AP003335
25	118.6	10.9	1131	6	AX449314
26	104.6	9.6	426	6	AX449520
27	98.8	9.1	1313	9	BC021113
28	98.8	9.1	1802	9	BC013993
29	98.8	9.1	2277	6	CQ717592
30	98.8	9.1	2320	6	AR380911
31	98.8	9.1	2320	9	HSDE2
32	97.6	9.0	1969	5	BC082841
33	97	8.9	1266	6	AX333694
34	97	8.9	1266	9	HUMDP2M
35	97	8.9	2968	6	CQ489116
36	97	8.9	2968	6	CQ494960
37	94.4	8.7	1902	5	BC080383
38	93.6	8.6	870	6	AX449515
39	93.6	8.6	1442	6	AX449342
40	88	8.1	1226	3	DMDPRTF
41	88	8.1	2200	3	AY028165
42	88	8.1	2203	6	CQ581589
43	88	8.1	2461	3	BT003183
44	88	8.1	2495	3	AY069526
45	86.6	8.0	2325	5	BC044409

## ALIGNMENTS

RESULT 1	AX100704	1089 bp	DNA	linear	PAT 10-APR-2001
LOCUS	Sequence 1 from Patent WO0121644.				
DEFINITION	AX100704				
ACCESSION	AX100704				
VERSION	AX100704.1	GI:13619652			
KEYWORDS	Triticum monococcum				
SOURCE	Triticum monococcum				
ORGANISM	Triticum monococcum				
REFERENCE	Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.				
AUTHORS	Wheat dp proteins and uses thereof				
TITLE	Patent: WO 0121644-A 1 29-MAR-2001;				
JOURNAL	CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)				
FEATURES	Location/Qualifiers				
source	1. .1089				
	/organism="Triticum monococcum"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4568"				
	20. .805				
	/notes="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC36466.1"				
	/db_xref="GI:13619653"				
	/translation="MAPPPGAAATAALDLTGVLHLEASSVPPLPERGNAVQRKG				
	AVDPDKRKKKAAAPRITGWLREYSKIVCEKAEKRTTNEVADEITYSELKMAH				
	IGQFDEKIRRRVDAFNLIALRVIAKEKEIRWGLSNRYEYKIKLEEVKELV				
	NKRNKALLOIEKQFDLONIKLNQTLSSAENVNGIRLPLVLKTSRKARVEIE				
	ISDSDSKFAHFEENGAPFTLHDDLSTILEGVRNRSIGRAGRTLH"				

## ORIGIN

Query Match	99.6%;	Score 1084.2;	DB 6;	Length 1089;
Best Local Similarity	99.7%;	Pred. No. 9.4e-229;		
Matches 1086;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0;			
QY	1	GAATTCGGACGAGCGCAATGCGCTCCCGCGGAGCTGCTGCGGCGCGCTACCGC	60	
Db	1	GAATTCGGACGAGCGCAATGCGCTCCCGCGGAGCTGCTGCGGCGCGCTACCGC	60	

```
QY 61 CGCAGTGGACCTGACCGCGGTGCACATTTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGA 120
Db 61 CGCAGTGGACCTGACCGCGGTGCACATTTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGA 120
QY 121 AGCGGGCGGTAAATCGCGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGAA 180
Db 121 AGCGGGCGGTAAATCGCGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGAA 180
QY 181 GGAGAAGGCTGGCGCACCGAGGATCACCGTTGGGGCTCCGGAGTACAGCAAAATAGT 240
Db 181 GGAGAAGGCTGGCGCACCGAGGATCACCGTTGGGGCTCCGGAGTACAGCAAAATAGT 240
QY 241 TTGTGAGAAAGTTGAAGCGCAAGCAAGCAACAATCAATGAGGTTCCAGACCAAAATTTA 300
Db 241 TTGTGAGAAAGTTGAAGCGCAAGCAAGCAACAATCAATGAGGTTCCAGACCAAAATTTA 300
QY 301 TTGAGAGCTGAAGTCCATGGCCACATATTTGTTCAAGGGTTTGTGAGAGGAATATTAGGCG 360
Db 301 TTGAGAGCTGAAGTCCATGGCCACATATTTGTTCAAGGGTTTGTGAGAGGAATATTAGGCG 360
QY 361 GAGAGTGTATGCTGTTTCAAGCTTCTCATTTGCACTTCGTTGTTATTGCAAAAGAAAAA 420
Db 361 GAGAGTGTATGCTGTTTCAAGCTTCTCATTTGCACTTCGTTGTTATTGCAAAAGAAAAA 420
QY 421 GGAGATACGCTGGATGGCGCTTTCAATTAACAGATATGAAAAATAAGAGCTTGAGGA 480
Db 421 GGAGATACGCTGGATGGCGCTTTCAATTAACAGATATGAAAAATAAGAGCTTGAGGA 480
QY 481 AGTTGCTAAAGAACTGCTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGA 540
Db 481 AGTTGCTAAAGAACTGCTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGA 540
QY 541 AAAACAGTTTGATGATCTCAAAACATCAAGTTACGTTAACCAACACTGGAAGCTCAGC 600
Db 541 AAAACAGTTTGATGATCTCAAAACATCAAGTTACGTTAACCAACACTGGAAGCTCAGC 600
QY 601 AGAGATGTTAATGGCATCGCCTTCCATTCGTTATGTTTGGTCAAGACATCTAGGAAGCAAG 660
Db 601 AGAGATGTTAATGGCATCGCCTTCCATTCGTTATGTTTGGTCAAGACATCTAGGAAGCAAG 660
QY 661 GGTGGAATTTGAGATTTGAGATGACTGCAAGATTTTGGCCATTTGCGAGTTCAATGGTGACC 720
Db 661 GGTGGAATTTGAGATTTGAGATGACTGCAAGATTTTGGCCATTTGCGAGTTCAATGGTGACC 720
QY 721 ATTACATTTGATGATGATCTCAATCTCTTGAAGGGGTAAAGGGTAAACGATAGGAAG 780
Db 721 ATTACATTTGATGATGATCTCAATCTCTTGAAGGGGTAAAGGGTAAACGATAGGAAG 780
QY 781 AGCTGGCGCGCCACCTTCACTAGACACTCAAGATATTACAAATGAAATTTAAAGTGT 840
Db 781 AGCTGGCGCGCCACCTTCACTAGACACTCAAGATATTACAAATGAAATTTAAAGTGT 840
QY 841 AGAAGTGGCAGCGCGGATTTTTCACAGCTATGATAGCTATATATCTCATGAAAA 900
Db 841 AGAAGTGGCAGCGCGGATTTTTCACAGCTATGATAGCTATATATCTCATGAAAA 900
QY 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGCTTTGAGAGATTTTAACTGCAAAATTTG 960
Db 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGCTTTGAGAGATTTTAACTGCAAAATTTG 960
QY 961 TCTCCTTTTGTGCTAGCAGGTTATTAGGCTTCAGATAGATGATTATATATGCTGTC 1020
Db 961 TCTCCTTTTGTGCTAGCAGGTTATTAGGCTTCAGATAGATGATTATATATGCTGTC 1020
QY 1021 TATGAAAAATTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1080
Db 1021 TATGAAAAATTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1080
QY 1081 AAAAAAAA 1089
Db 1081 AAAAAAAA 1089
```

```
RESULT 2
TSP271917 1083 bp mRNA linear PLN 16-DEC-2000
LOCUS Triticum sp. mRNA for DP protein (dp gene).
DEFINITION AJ271917
ACCESSION AJ271917
VERSION GI:11877790
KEYWORDS DP gene; E2F dimerization partner.
SOURCE Triticum sp.
ORGANISM Triticum sp.
REFERENCE 1
AUTHORS Ramirez-Parra, R. and Gutierrez, C.
TITLE Characterization of wheat DP, a heterodimerization partner of the
JOURNAL Plant E2F transcription factor which stimulates E2F-DNA binding
MEDLINE FEBS Lett. 486 (1), 73-78 (2000)
PUBMED 20562804
REFERENCE 2 (bases 1 to 1083)
AUTHORS Gutierrez, C.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Gutierrez C., C.S.I.C.-U.A.M., Centro de
Biologia Molecular, Cantoblanco, Madrid 28049, SPAIN
FEATURES
source
1..1083
/mol_type="mRNA"
/db_xref="taxon:4569"
1..1083
/gene="dp"
<1..13
/gene="dp"
14..799
/gene="dp"
/function="cell cycle E2F dimerization partner"
/codon_start=1
/product="DP protein"
/protein_id="CAC19034.1"
/db_xref="GI:11877791"
/db_xref="GOA:Q9FET1"
/db_xref="UniProt/TrEMBL:Q9FET1"
/translat="MAPPRGAAAAAALDLTGVLHLEASSVPPLPERGNAVORRG
AVDPDRKKEKAAAPRTGWLREYSKIVCEKVEAKGRTTYNEADEIYSELSMAH
IGGPFDEKIRRVDAFNVLIALRVIAKEKEIRWGLSNRYEYKIKLEVRKELV
NKIRKKALLQEIEKOFDDLQNKLRNQLTESSAENVNGIRLPFLVLTGSKARVEIE
ISDTSKFAHFEFNGAPFTLHDDLILLEGVRGNSIGKAGRATLH"
801..1033
/gene="dp"
3'UTR
ORIGIN
Query Match 97.8%; Score 1065.4; DB 8; Length 1083;
Best Local Similarity 99.0%; Pred. No. 1.3e-224;
Matches 1072; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 7 GGACAGACCGCAATGGCGCTCCCGCGCGAGCTGTGCGCGCGCTACCGCGCACT 66
Db 1 GGACAGACCGCAATGGCGCTCCCGCGCGAGCTGTGCGCGCGCTACCGCGCACT 60
QY 67 GGACCTGACCGCGGTGCACATTTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGAAGCGGG 126
Db 61 GGACCTGACCGCGGTGCACATTTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGAAGCGGG 120
QY 127 CGGTAAATGCGGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGGAAGAA 186
Db 121 CGGTAAATGCGGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGGAAGAA 180
QY 187 GGCTGGGCAACCGAGGATCACCGGTGGGGCTCCCGAGTACAGCAAAATAGTTTGA 246
Db 181 GGCTGGGCAACCGAGGATCACCGGTGGGGCTCCCGAGTACAGCAAAATAGTTTGA 240
QY 247 GAAAGTTGAAGCCAAAGGAAGAAACAATACATAGAGTTGCAGACGAAATTTATTTCAGA 306
```

Db	241	GAAGTTGAAGCCAAAGAGAAACAACATCAATGAGGTTCAGACGCAATTTATTTCAGA	300
Qy	307	GCTGAAGTCCATGGCACAATATTTGCTCAAGGTTTGATGAGAAGAAATATTAGGCGGAGAT	366
Db	301	GCTGAAGTCCATGGCACAATATTTGCTCAAGGTTTGATGAGAAGAAATATTAGGCGGAGAT	360
Qy	367	GTATGATGCTTTCAACGTTCTCATTCGACTTCGTTGTTATTGCAAAAGAAAAAGGAGAT	426
Db	361	GTATGATGCTTTCAACGTTCTCATTCGACTTCGTTGTTATTGCAAAAGAAAAAGGAGAT	420
Qy	427	ACGCTGATGGCCCTTTCAATATTCAGATATGAAAAATTAAGAGCTTCGAGGAGTTTCG	486
Db	421	ACGCTGATGGCCCTTTCAATATTCAGATATGAAAAATTAAGAGCTTCGAGGAGTTTCG	480
Qy	487	TAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGCAAAATCGAAAAACA	546
Db	481	TAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGCAAAATCGAAAAACA	540
Qy	547	GTTTGATGATCTCCAAACATCAAGTTACGTTACGTAACCAACACTCGAAAGCTCAGCAGAGAA	606
Db	541	GTTTGATGATCTCCAAACATCAAGTTACGTTACGTAACCAACACTCGAAAGCTCAGCAGAGAA	600
Qy	607	TGTTAATGGCATCGCCCTTCATTCGTTATGCTCAAGCATCTAGGAAGCAAGGTTGGA	666
Db	601	TGTTAATGGCATCGCCCTTCATTCGTTATGCTCAAGCATCTAGGAAGCAAGGTTGGA	660
Qy	667	AATTGAGATTTTCAGATCACTCGAAGTTTGCCATTTTCGAGTTCAATGCTGACCATTCAC	726
Db	661	AATTGAGATTTTCAGATCACTCGAAGTTTGCCATTTTCGAGTTCAATGCTGACCATTCAC	720
Qy	727	ATTGTCATGATGATCTCTCAATCTTTGAGGGGGTAAGCGGTAACAGCATAGGAAGCTGG	786
Db	721	ATTGTCATGATGATCTCTCAATCTTTGAGGGGGTAAGCGGTAACAGCATAGGAAGCTGG	780
Qy	787	CCGGCCACCCCTTCATGAGACTCAAGATATTTACAAATGAATTAAGTTGTTAGAACT	846
Db	781	CCGGCCACCCCTTCATGAGACTCAAGATATTTACAAATGAATTAAGTTGTTAGAACT	840
Qy	847	GGCACAGCCGATCTTTTCACAGCTATGATAGCTATATATCTCATGAAACTTGAC	906
Db	841	GGCACAGCCGATCTTTTCACAGCTATGATAGCTATATATCTCATGAAACTTGAC	900
Qy	907	CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGATTTTAACTGCAAAATTTGTCCT	966
Db	901	CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGATTTTAACTGCAAAATTTGTCCT	960
Qy	967	TTTTGTCCTAGCAGGTTTATGCTTCAGATAGATTCATATATGCTGCTATGAA	1026
Db	961	TTTTGTCCTAGCAGGTTTATGCTTCAGATAGATTCATATATGCTGCTATGAA	1020
Qy	1027	AACATTGATGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1086
Db	1021	ACCATGATGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1080
Qy	1087	AAA 1089	
Db	1081	AAA 1083	
LOCUS	AK111611	1174 bp mRNA linear	PLN 29-OCT-2003
DEFINITION	Oryza sativa (japonica cultivar-group)	CDNA clone:J013105118, full insert sequence.	
ACCESSION	AK111611		
VERSION	AK111611.1	GI:37988274	
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1		

## AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, E., Sugano, S., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Iwata, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayase, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

## TITLE

Japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

## JOURNAL

Japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

## MEDLINE

Japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

## PUBMED

Japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

## AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayase, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shiehiki, T., Sugabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tanaka, T., Tomaru, A., Toyata, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

## Rice full-length cDNA

Rice full-length cDNA

## Unpublished

Unpublished

## 3 (bases 1 to 1174)

3 (bases 1 to 1174)

## Kikuchi, S.

Kikuchi, S.

## Direct Submission

Direct Submission

Submitted (12-SEP-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)

Tel: 81-29-838-7007, Fax: 81-29-838-7007

This clone is one of the 32K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, E., Sugano, S., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Iwata, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayase, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice.

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayase, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shiehiki, T., Sugabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tanaka, T., Tomaru, A., Toyata, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

## TITLE

Rice full-length cDNA

## JOURNAL

Rice full-length cDNA

## REFERENCE

Unpublished

## AUTHORS

Kikuchi, S.

## TITLE

Direct Submission

## JOURNAL

Submitted (12-SEP-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp)

## COMMENT

This clone is one of the 32K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, E., Sugano, S., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Iwata, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayase, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice.

Science 301 (5631), 376-379 (2003)

22752273

12863764

2

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayase, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shiehiki, T., Sugabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tanaka, T., Tomaru, A., Toyata, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

FEATURES		source
1. 1174 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="J013105118"		
ORIGIN		
Query Match	47.3%; Score 515; DB 8; Length 1174;	
Best Local Similarity	77.4%; Pred. No. 4.1e-103;	
Matches	722; Conservative 0; Mismatches 185; Indels 26; Gaps 7;	
Qy	118 CGAAGCGGCGGTAAATCGGTTCCAAAGGAAGGGGCTGTTGACCCGGATAAAGATAGAA	177
Db	217 CGGCGGCGGCAATCGGCAGAGAGGGAGGGCGGTG--CGAACAGGAACGGAA	273
Qy	178 GAAGGAAGGCTCGGCACCGAGGNTCACGGTTGGGGGCTCGCGAGTACAGCAAAAT	237
Db	274 GAAGGAAGACCGGGCGCAGAGGATCACCGGTGGGGGCTCGTGAGTTTCGCAAGAT	333
Qy	238 AGTTTGTGAGAAAGTTCGAAGCCAAAGGAAGAACACATACAAATCAGAGTTGCAGACGAAT	297
Db	334 AGTTTCTAAGAAAGTTGAGGCCAAGGAGAACACATATATCAGGTTGCCGATGAGAT	393
Qy	298 TTATTCAGAGCTGAAGTCCATGCGCACATATTTGTTCAAGGGTTTGATGAGAAGATATTAG	357
Db	394 TTTTTCGGAGCTGAAGTCCATTAAGCAGAACCGGTCGGAGTTTGATGAGAAGAAATATTAG	453
Qy	358 GCGGAGAGTGTATCATGCTTTCAAGTTCTCATTTGCACTTCGTGTTATTGCAAAAGAAA	417
Db	454 GCGGAGGATATATGATGCTTTCAATGTGCTCATTTGCAATTCGTGTTATTGCAAAAGTAA	513
Qy	418 AAAGGAGATACGGTGGATGGGCCCTTTCAAATTTACAGATATGAAAAAATAAAGAGCTTGA	477
Db	514 AAAGAGATAAAGTGGATGGGCCCTTACTAATATTAGATACGAAAAGATACAGAAGTTGGA	573
Qy	478 GGAAGTTGTTAAGAACTCGTTCACACAGATTAGGACCAAGAGGCACCTCCTCCAGGAAT	537
Db	574 GGAAAGTTCACAAAAGACTCATCACAGGATCAAGATCAAGAAAGAGCTTCTCCAGGAAT	633
Qy	538 CGAAAAACAGTTTGATGATCTCCAAAAACATCAAGTTACGTAAACCAACACCTCGAAAGCTC	597
Db	634 TGAAGAGCAGTTTCGATGACCTTCAGATATTACATTACGCAACGAGGCTAGTCAGAGGCC	693
Qy	598 AGCAGAGAAATGTTAATGGCATCCGGCTTCATTCGTAATTTGTTGTCAGACATCTAGGAAGC	657
Db	694 AGCAGAAAGTGTAAATGGCATCCTCCTTCGTTCTTATTGATCAAGACATCCCGAAAAAGC	753
Qy	658 AAGGTTGGAATTCGAGATTTTCAGATGACTTCGAAGTTTCCCATTTTCGAGTTTCAATGGTGC	717
Db	754 AAGGTTGGAATTTGAGATTTTCGGAAGATTCGAAGTTTGCACGGTTTCGCTTCACGGTGC	813
Qy	718 ACCATTACATTTGCATGATGATCTCTCAATCTTTGAGGGGGTAAAGCGGTAAACAGCATAGG	777
Db	814 ACCATTACCATGATGATGATGATCAATCCTTGAAGCCATCAGGCGTAA---CAAAAGG	870
Qy	778 AAGAGCTGGCGCGCCACCTTCACT-AGAGACTCAAGATATTACAAATGAAATGAAATGAAAG	836
Db	871 AAGAGCTGGCCCTCTCCATTACCTTTAAGAGGCAACAGATGTGCAAAACCAATTGGAAG	930
Qy	837 TG-----TTAGAACTGGCACGCGGATTCCTTTTGACACAGCTATGTATAGCTATA---886	



```

KEYWORDS      FLJ: CDNA.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS        Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                1 (bases 1 to 879)
                Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
                Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,
                Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
                Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
                Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
                Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
                Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and
                Ecker, J.R.
TITLE          Arabidopsis ORF clones
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 879)
AUTHORS        Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
                Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Huan, V.W.,
                Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
                Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
                Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
                Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
                Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and
                Ecker, J.R.
TITLE          Direct Submission
JOURNAL        Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory
                (SIGnAL), Plant Biology Laboratory, The Salk Institute for
                Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
                USA
COMMENT        RIKEN Genomic Sciences Center (GSC) members carried out the
                collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
                Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
                Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
                Hayashizaki, Y. and Shinzaki, K.
                The Salk, Stanford, PESC (SSP) Consortium members constructed and
                sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J.,
                Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
                Dale, J.M., Huan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
                Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
                Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
                Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
                Ecker, J.R.
                Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
                this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
                contributed equally to this work as PIs.
FEATURES      Location/Qualifiers
                1..879
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /chromosome="5"
                /clone="U60005"
                /ecotype="Columbia"
                /note="This clone is in pUNI 51"
                1..879
                /code="putative DP-2 transcription factor"
                /codon_start=1
                /product="At5g02470"
                /protein_id="AA063350.1"
                /db_xref="GI:28950853"
                /translation="MSMEMELFVPEKQHPVSVEKTPVRRLIVDDSEIGSEKK
                QGSTSGGLRQFVSMVCKLEKITYKEVADEIISDFATIKONAEKPLNNEVNE
                KNIRRVYDALNFMALDIARDKKEIRWGLPTCKDKDVEKMDKNVSSVOKKA
                AFLKELREKVSLESLSNRQEMVVKQCPAEGFTLFLILENPHAVIEISDMQ
                LVHLDFSTPFSVHDDAYILKLMQEQKQNRVSSSSSTTHQSHSSAHSSSSCIAS
                GTSGPVCNWSGSDTR"
CDS           14.1k; Score 153.6; DB 8; Length 879;
ORIGIN
Query Match
14.1k; Score 153.6; DB 8; Length 879;

```

```

Best Local Similarity 55.6%; Pred. No. 2.3e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;
QY 131 AATCGCGTCCAAAGGAAGGGCGCTGTTGACCCGGATAAAGATAGGAAGAGGAGCT 190
DB 73 ACTCCAGTGAGAAGAAATGATTGTTGATGATGATCTCGAAATGGATCAGAGAAGAA 132
QY 191 GCGGACCGAGGATCACCGGTTGGGGCTCCGGAGTACAGCAAAATAGTTGTGAGAAA 250
DB 133 GGGCAATACAGAACTTCTGGAGCGGGCTTCGTCATTCAGTGTATGTTGTTGTCAGAAG 192
QY 251 GTTGAAGCCAAAGGAAGAACACATACATAGAGGTTGCAGACGAAATTTATTCAGAGCTG 310
DB 193 TTGAAGCCCAAGAAAGATACTTTACAAGAGGTTGCAGACGAAATTTATTCAGATTTT 252
QY 311 AAGTCCAT-----GGCACAATATTGGTCAAGGGTTTGTATGAGAAGAA 352
DB 253 GCCACAATTAAGCAAAACGAGAGAAGCCTTTCAATGAAATGAGTACATGAGAAGAAC 312
QY 353 ATTAGCGGAGAGTGATGATGCTTTCAAGCTTCTCATTCGACCTTCGTGTTATTCAGAAA 412
DB 313 ATAAGCGGAGAGTCTACGATCGGCTCAATGTGTTTCATGCGCTTGGATATTTATTCGAAG 372
QY 413 GAAAAAAGGAGATACGTTGATGGGCTTTTCAAAATTTACAGATATGAAAAATAAAGAAG 472
DB 373 GATAAAAAGGAATCCGTGGAAAGACTTCCTATTACCTGCAAAAGGATGGGAAGAA 432
QY 473 CTTGAGAGAGTTCGTAAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCTCCAG 532
DB 433 GTCAAGATGATCGTAATAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTCTTAAA 492
QY 533 GAAATCGAAAAACAGTTTGATGATCTCCAAAACATCAAGTTAGTACGAA---ACACTG 589
DB 493 GAGTTGAGAGAAAAGGCTCAAGCTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 552
QY 590 GAAAGCTCAGCAGAGAATGTAATGGCATCCGCTTCATTCGTTATGTTGGTCAAGACATCT 649
DB 553 GTGAAGACTCAAGCCAGCAGAGAGATTACCTTACCATTCATCTTACTTGAGAGAAC 612
QY 650 AGGAAAGCAAGGGTGGAAATTTAGATTTTCAGATGACTCGAAGTTTGCCTTCGAGTTC 709
DB 613 CCTCAGCAGTAGTCGAAATCGAGATTTCTGAAGATATGCAACTTGTACACCTCGACTTC 672
QY 710 AATGGTGCCACCATTCACATTCGATGATGAT 739
DB 673 AATAGCACACCTTTCTCGGTCCATGATGAT 702
RESULT 8
ATH319027
LOCUS         Arabidopsis thaliana partial mRNA for E2F dimerisation partner
DEFINITION   protein (dp2a gene).
ACCESSION    AJ319027
VERSION      AJ319027.1 GI:18447785
KEYWORDS     qp2a gene; E2F dimerisation partner protein.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1
AUTHORS      Rossignol, P.
TITLE        E2F family transcription factors: AtE2F-a and AtDP-a, induce
              Arabidopsis leaf cells to re-enter S phase
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 960)
AUTHORS      Bergounioux, C.
TITLE        Direct Submission
JOURNAL      Submitted (29-AUG-2001) Bergounioux C., UMR 8618, Cnrs, IBP bat 630
              Université Paris-Sud, 91405 Orsay, FRANCE
FEATURES     Location/Qualifiers
              1..960

```



```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Colombia"
/db_xref="taxon:3702"
/tissue_type="buds"
/dev_stage="flowering"
1..960
/gene="dp2a"
<1..>960
/gene="dp2a"
79..957
/gene="dp2a"
/function="transcription"
/codon_start=1
/product="E2F dimerisation partner protein (DP2a)"
/protein_id="CAC87459.1"
/db_xref="GI:18447786"
/db_xref="GOA:Q9FNY3"
/db_xref="UniProt/T/EMBL:Q9FNY3"
/translations="MSMEMELFTVPEKORHPSVSVKTPVRRLIVDDSEIGSEKK
GQSRISGGIRQPSVWCQKLEAKITTYKEVADEILSDPATIKQNAKPKLNENEYNE
AFIKRRVYDALNVFMALDIITARDKKEIRWGLPITCKDVEEVMKDRNKVMYSQVKKKA
NKLREKYSLSLEMSRQSEVMVYTKGPAEGFTLPFILLETNPFAVVEIIESDQKA
LVHLDFNSTPFSVDDHDAIYLKLMQEQKQNRVSSSSSTHHQSHSSAHSSSSSCIAS
GTSCPVCNWSGSDINTP"

```

Best Local Similarity 55.6%; Pred. No. 2.4e-23;  
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

	Qy	131	AATCGCGTCCAAAGGAAGGGGGCTGTTTGACCGGATAAGAGATAGGAAGAAGGAGAAGGCT	190
	Db	86	ACTCCAGTGAGAAGGAAATTGATTGTTGATGATGATTCTGAAATTGGATCAGAGAAGAAA	145
	Qy	191	GCGGCCAGGAGATCACCGGTTGGGGCTCGCGGAGTACAGCAAAATAGTTTGTGAGAAA	250
	Db	146	GGGCAATCAAGAAGCTCTCGAGGCGGGCTTCGTCAATTCAGTGTATTGGTTTTTCAGAAG	205
	Qy	251	GTTGAAGCCAAAGGAAGAACACATACAATGAGGTTCGACGCAAAATTTATTTCAGAGCTG	310
	Db	206	TTGGAAGCCAAAGAAGATAACTACTTACAAGGAGGTTCGACGCAAAATTTATTTCAGATTTT	265
	Qy	311	AAGTCCAT-----GGCACATATTGGTCAAGGGTTTGATGAGAAGNAT	352
	Db	266	GCACAAATTAAACAAACCGCAGAGAAGCCTTTGAATGAAATGAGTACAATGAGAAGAAC	325
	Qy	353	ATTAGGCGGAGAGTGTATGATGCTTTTCAACGCTTCTCATTTGCACATTCGTGTTATTTCGAAAA	412
	Db	326	ATAAGGCGGAGAGTCTACGATGCGTCAATGTGTTCATGGCGTTGGGATATTTATTTCGAAGG	385
	Qy	413	GAATAAAGGAGATACGGTCGATGGGCTTTCAAAATTACAGATATGAATAAATAAAGNAG	472
	Db	386	GATAAAAGGAAATCCGGTGGAAAGGACTTCCCTATTACCTGCAAAAAGGATGTGGAAAGAA	445
	Qy	473	CTTGAGGAAGTTCGTAAAGAACTCGTCAACAAGATTTAGGAACCAAGAAGGCACCTCCTCCAG	532
	Db	446	GTCAAGATGCGTAAATAAGTTATGAGCAGTGTGCAAAAGAAGCGTCTTTTCTTTAAA	505
	Qy	533	GAATTCGAAAAACAGTTTGTATGATCTCCAAAAATCAAGTTACGTAAACCAA--ACACTG	589
	Db	506	GAGTTGAGAAAGGCTTCAAGTCTTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT	565
	Qy	590	GAAAGCTCAGCAGAGAATGTTAATGGCATCCGCGCTTCCATTTCGTTATTGGTCAAGACATCT	649
	Db	566	GTGAAGACTCAAGGCCCAGCAGAAGGATTTACCTTACCATTTCTTACTTGAGACAAAC	625
	Qy	650	AGGAAGCAAGGCTGGAAATTGAGATTTCAAGTACTCGAAGTTGGCCCATTTTCAGTTTC	709
	Db	626	CCTCAGCAGTAGTGCAGAAATTCGAGATTTCTGGAAGATATGCAACTTGTACACCTCGATTC	685
	Qy	710	AATGGTGCACCAATTCACATTCGATGATGAT	739



Db 161 ACTCCAGTGAGAGAAATTAATGTTTGTATGATGATCTTGAATTTGGATCAGAGAGAAA 220  
Qy 191 GCGGCACCGAGGATCACCGGTTGGGGCTCCGGGAGTACAGCAAAATAGTTTGTGAGAAA 250  
Db 221 GGGCAATCAAGAACTTCTGAGGCGGCTTCGTCATTTAGTGTATGTTTGTGAGAG 280  
Qy 251 GTTGAAGCCAAAGAGAAACAACATACAAATGAGGTTCCAGACGAAATTTATTCAGAGCTG 310  
Db 281 TTGGAAGCCAAAGAGATAACTACTTACAGAGGTTCCAGACGAAATTTATTCAGATTTT 340  
Qy 311 AAGTCCAT-----GGCACATATTGGTCAAGGTTTGTGAGAGAGAAAT 352  
Db 341 GCCACAATTAAGCAAAACGAGAGAGCCCTTGAATGAAATGAGTACAAATGAGAGAAC 400  
Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTCGACTTCGTTGTTATTCAGAAA 412  
Db 401 ATAAGCGGAGAGTCTAGATGCGCTCAATGTTTCAATGCGGTTGATATTTATTCAGAG 450  
Qy 413 GAAAAAAGAGAGATACGTTGGATGGGCTTTTCAAAATPACAGATATGAAAAATAAAGAG 472  
Db 461 GATAAAAAGGAAATCCGGTGGAAAGGACTTCTTATTACCTGCAAAAGAGGATGTGAAGAA 520  
Qy 473 CTTGAGGAAGTTGTAAGAACTCGTCAACAGATTTAGGAACAGAGAGGCACTCTCCAG 532  
Db 521 GTCAAGATGATCGTATAAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAAA 580  
Qy 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTACGTAAACAA---ACACTG 589  
Db 581 GAGTTGAGAGAAAGGTTCTCAAGTCTTGAGAGCTTATGTCGAGAAATCAAGAGATGGTT 640  
Qy 590 GAAAGCTCAGCAGAGAAATGTAATGGCATCGGCTTCCATTCTGATTTGGTCAAGACATCT 649  
Db 641 GTGAAGACTCAAGCCCGCAGAGAGGATTTACCTTACCATTCATCTACTTTGAGACAAAC 700  
Qy 650 AGGAAGCAAGGTTGGAATTTGAGATTTCAAGATTCAGATGATCGAAGTTTGCCATTTGAGATTC 709  
Db 701 CTTCAAGCAGTAGTCGAAATTCGAGATTTCTGAAGATATGCAACTTTGACACCTCGACTTC 760  
Qy 710 AATGGTGCACCATTCACATTCGATGATGAT 739  
Db 761 AATAGCACACCTTTCTCGTCCATGATGAT 790

RESULT 12  
AX712207  
LOCUS AX712207 1274 bp DNA linear PAT 11-APR-2003  
DEFINITION Sequence 3 from Patent WO03018818.  
ACCESSION AX712207  
VERSION AX712207.1 GI:29823429  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE de Veylder L., Inze, D., Mironov, V. and Segers, G.  
AUTHORS Method and means for modulating plant cell cycle proteins and their  
TITLE use in plant cell growth control  
JOURNAL Patent: WO 03018818-A 3 06-MAR-2003;  
CropDesign N.V. (BE)  
FEATURES Location/Qualifiers  
source 1..1274  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 14.1%; Score 153.6; DB 6; Length 1274;  
Best Local Similarity 55.8%; Pred. No. 2.4e-23;  
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;  
Qy 131 AATGCGGTCCAAAGGAAGGGGCTGTTGACCGGATAAAGATAGGAAGGAGGAGGCT 190

Db 246 ACTCCAGTGAGAGAAATTAATGTTTGTATGATGATCTTGAATTTGGATCAGAGAGAAA 305  
Qy 191 GCGGCACCGAGGATCACCGGTTGGGGCTCCGGGAGTACAGCAAAATAGTTTGTGAGAAA 250  
Db 306 GGGCAATCAAGAACTTCTGAGGCGGCTTCGTCATTTAGTGTATGTTTGTGAGAG 365  
Qy 251 GTTGAAGCCAAAGAGAAACAACATACAAATGAGGTTTCAGACGAAATTTATTCAGAGCTG 310  
Db 366 TTGGAAGCCAAAGAGATAACTACTTACAGAGGTTTCAGACGAAATTTATTCAGATTTT 425  
Qy 311 AAGTCCAT-----GGCACATATTGGTCAAGGTTTGTGAGAGAGAAAT 352  
Db 426 GCCACAATTAAGCAAAACGAGAGAGCCCTTGAATGAAATGAGTACAAATGAGAGAAC 485  
Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTCGACTTCGTTGTTATTCAGAAA 412  
Db 486 ATAAGCGGAGAGTCTACGATGCGCTCAATGTTTCAATGCGGTTGATATTTATTCAGAG 545  
Qy 413 GAAAAAAGAGAGATACGTTGGATGGGCTTTTCAAAATPACAGATATGAAAAATAAAGAG 472  
Db 546 GATAAAAAGGAAATCCGGTGGAAAGGACTTCTTATTACCTGCAAAAGAGGATGTGAAGAA 605  
Qy 473 CTTGAGGAAGTTGTAAGAACTCGTCAACAGATTTAGGAACAGAGAGGCACTCTCCAG 532  
Db 606 GTCAAGATGATCGTATAAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAAA 665  
Qy 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTACGTAAACAA---ACACTG 589  
Db 666 GAGTTGAGAGAAAGGTTCTCAAGTCTTGAGAGCTTATGTCGAGAAATCAAGAGATGGTT 725  
Qy 590 GAAAGCTCAGCAGAGAAATGTAATGGCATCGGCTTCCATTCTGATTTGGTCAAGACATCT 649  
Db 726 GTGAAGACTCAAGCCCGCAGAGAGGATTTACCTTACCATTCATCTACTTTGAGACAAAC 785  
Qy 650 AGGAAGCAAGGTTGGAATTTGAGATTTCAAGATGATCGAAGTTTGCCATTTGAGATTC 709  
Db 786 CTTCAAGCAGTAGTCGAAATTCGAGATTTCTGAAGATATGCAACTTTGACACCTCGACTTC 845  
Qy 710 AATGGTGCACCATTCACATTCGATGATGAT 739  
Db 846 AATAGCACACCTTTCTCGTCCATGATGAT 875

RESULT 13  
AX734251  
LOCUS AX734251 1274 bp DNA linear PAT 08-MAY-2003  
DEFINITION Sequence 3 from Patent WO03025185.  
ACCESSION AX734251  
VERSION AX734251.1 GI:30513582  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE Beeckman, T., de Veylder, L., Inze, D., Mironov, V., Broekaert, W.,  
AUTHORS Dillen, W. and Frankard, V.  
TITLE A method to modify cell number, architecture and yield of plants by  
overexpressing the e2f transcription factor  
JOURNAL Patent: WO 03025185-A 3 27-MAR-2003;  
CropDesign N.V. (BE)  
FEATURES Location/Qualifiers  
source 1..1274  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 14.1%; Score 153.6; DB 6; Length 1274;  
Best Local Similarity 55.6%; Pred. No. 2.4e-23;  
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

```
Qy 131 AATCGGTCCAAAGAGAGGGGCTGTTGA CCCGGATAAAGATAGGAAGAGGAGGCT 190
Db 246 ACTCCAGTGAAGAGAAATGATTGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 305

Qy 191 GCGCACCGAGGATCACCGGTTGGGGCTCGCGAGTACAGCAAAATAGTTGTGAGAA 250
Db 306 GGGCAATCAGAACTTCTGGAGCGGGCTTCGTCAATTCTAGTGTATGGTTGTGAGAAG 365

Qy 251 GTTGAAGCCAAAGGAAGAACACATACAATGAGTTGCAGACGAAATTTATTCAGAGCTG 310
Db 366 TTGGAAGCCAAAGGAAGAACATACAATGAGTTGCAGACGAAATTTATTCAGATTT 425

Qy 311 AAGTCCAT-----GGCACAATTTGGTCAAGGGTTGATGAGAGAAAT 352
Db 426 GCCACAATTTAAGCAAAACGAGAGAGGCTTTGAATGAAATGATGATAATGAGAGAAC 485

Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTTGCACCTCGCTGTTATTCAGAA 412
Db 486 ATAAGCGGAGAGTCTAGATCGCTCAATGTGTTCAATGCGGTTGGATATATTTGCAAG 545

Qy 413 GAAAAAAGGAGATACCGGTGGATGGGCTTTCAAATTTACAGATATGAAAAATAAAGAG 472
Db 546 GATAAAAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAAGGATGTGGAAGAA 605

Qy 473 CTTGAGGAGTTCGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCTCCAG 532
Db 606 GTCAAGATGATCGTAATAAGTTATGAGCAGGTGCAAAAGAGGCTGCTTTCTTAAA 665

Qy 533 GAAATCGAAAAACAGTTGATGCTCAAAACATCAAGTTTACGATGCTGAGATTGCCATT 589
Db 726 GTGAAGACTCAAGCCAGAGAGGATTTACCTTACCATTCTTCTTGTGAGACAAAC 785

Qy 590 GAAAGCCAGAGTTCGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCTCCAG 649
Db 725 GAGTTGAGAGAAAGGCTTCAAGTCTTGAGAGTCTTATGTGAGAGAAATCAAGAGATG 725

Qy 590 GAAAGCTCAGCAGAGAAATGTAATGATCGGCTTCCATTCTGTTGTTGTTGTTGTTG 649
Db 726 GTGAAGACTCAAGCCAGAGAGGATTTACCTTACCATTCTTCTTGTGAGACAAAC 785

Qy 650 AGGAAGCCAGGTTGGAATGATGATTCAGATGCTGAGATTGCCATTTCGAGTTTC 709
Db 786 CCTCAGCAGTAGTCGAAATCGAGATTTCTGAAGATATGCAACTTTGACACCTCGACT 845

Qy 710 AATGGTGACCATTCACATTCGATGATGAT 739
Db 846 AATAGCACACTTCTCGGTCCATGATGAT 875
```

```
RESULT 14
LOCUS   ATH294531               1274 bp    mRNA    linear    PLN 15-DEC-2000
DEFINITION   Arabidopsis thaliana mRNA for DP-like protein (dpa gene).
ACCESSION   AJ294531
VERSION     AJ294531.1   GI:11125650
KEYWORDS    DP-like protein; dpa gene.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Magyar, Z., Atanassova, A., De Veylder, L., Rombauts, S. and Inze, D.
Characterization of two distinct DP-related genes from Arabidopsis
thaliana
PFBs Lett. 486 (1), 79-87 (2000)
20562805
MEDLINE
PUBMED 11108847
REFERENCE 2 (bases 1 to 1274)
AUTHORS Magyar, Z.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2000) Magyar Z., Department Plantgenetica,
Laboratorium voor Genetica, Universiteit Gent, K.L. Ledeganckstraat
35., B-9000, Gent, BELGIUM
Location/Qualifiers
1. .1274
source
```

```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="Columbia"
/db_xref="taxon:3702"
1. .1274
/genes="dpa"
174. .1052
/genes="dpa"
/EC number="6.3.2.19"
/function="cell cycle regulation"
/codon_start=1
/product="DP-like protein"
/protein_id="CAC15483.1"
/db_xref="GI:11125651"
/db_xref="GOA:Q9FNY3"
/db_xref="Uniprot/TREMBL:Q9FNY3"
/translators="MSMEMELFVPEKQHPVSVEKTPVRKLIIVDDSEITGSEKK
GOSRTSGGLRQFVSVVCKLEAKITTYKEVADEIISDPATIKQAEKPLNENYKE
KNIRRVYDALNVFMAIDIIARDKEIRKGLPIITCKQVVEVEMKRNKVMSSVQKKA
APLKEKRVSSLESLSMRNQEMVVKITQGPAGFTIPLFILLETHPHAVVIEISEDMQ
LVHLDNFSTPFSVHDDAYILKMQEQBQBNRVSSSSTTHQSHSSSSSSCIAS
GTSGPVCMNSGSDTR"
```

## ORIGIN

```
Query Match 14.1%; Score 153.6; DB 8; Length 1274;
Best local Similarity 55.8%; Pred. No. 2.4e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

Qy 131 AATCGGTCCAAAGAGAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGGAGGCT 190
Db 246 ACTCCAGTGAAGAGAAATGATTGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 305

Qy 191 GCGCACCGAGGATCACCGGTTGGGGCTCGCGAGTACAGCAAAATAGTTGTGAGAA 250
Db 306 GGGCAATCAGAACTTCTGGAGCGGGCTTCGTCAATTTCAAGTGTATGGTTGTGAGAAG 365

Qy 251 GTTGAAGCCAAAGGAAGAACACATACAATGAGTTGCAGACGAAATTTATTCAGAGCTG 310
Db 366 TTGGAAGCCAAAGGAAGAACATACAATGAGTTGCAGACGAAATTTATTCAGATTT 425

Qy 311 AAGTCCAT-----GGCACAATTTGGTCAAGGGTTGATGAGAGAAAT 352
Db 426 GCCACAATTTAAGCAAAACGAGAGAGGCTTTGAATGAAATGATGATAATGAGAGAAC 485

Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTTGCACCTCGCTGTTATTCAGAA 412
Db 486 ATAAGCGGAGAGTCTAGATCGCTCAATGTGTTCAATGCGGTTGGATATATTTGCAAG 545

Qy 413 GAAAAAAGGAGATACCGGTGGATGGGCTTTCAAATTTACAGATATGAAAAATAAAGAG 472
Db 546 GATAAAAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAAGGATGTGGAAGAA 605

Qy 473 CTTGAGGAGTTCGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCTCCAG 532
Db 606 GTCAAGATGATCGTAATAAGTTATGAGCAGGTGCAAAAGAGGCTGCTTTCTTAAA 665

Qy 533 GAAATCGAAAAACAGTTGATGCTCAAAACATCAAGTTTACGATGCTGAGATTGCCATT 589
Db 666 GAGTTGAGAGAAAGGCTTCAAGTCTTGAGAGTCTTATGTGAGAGAAATCAAGAGATG 725

Qy 590 GAAAGCTCAGCAGAGAAATGTAATGATCGGCTTCCATTCTGTTGTTGTTGTTGTTG 649
Db 726 GTGAAGACTCAAGCCAGAGAGGATTTACCTTACCATTCTTCTTGTGAGACAAAC 785

Qy 650 AGGAAGCCAGGTTGGAATGATGATTCAGATGCTGAGATTGCCATTTCGAGTTTC 709
Db 786 CCTCAGCAGTAGTCGAAATCGAGATTTCTGAAGATATGCAACTTTGACACCTCGACT 845

Qy 710 AATGGTGACCATTCACATTCGATGATGAT 739
Db 846 AATAGCACACTTCTCGGTCCATGATGAT 875
```

```

RESULT 15
AX100710      153 bp      DNA      Linear      PAT 10-APR-2001
LOCUS
DEFINITION   Sequence 7 from Patent WO0121644.
ACCESSION   AX100710
VERSION     AX100710.1 GI:13619658
KEYWORDS
SOURCE      Triticum monococcum
ORGANISM    Triticum monococcum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
AUTHORS     Gutierrez-Amenta, C.C. and Ramirez-Parra, E.C.
TITLE       Wheat dp proteins and uses thereof
JOURNAL     Patent: WO 0121644-A 7 29-MAR-2001;
            CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source      1..153
            /organism="Triticum monococcum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4568"
            <1..>153
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC36469.1"
            /db_xref="GI:13619659"
            /translation="DKDRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADE
            IYSELKS"

CDS
            153

ORIGIN
Query Match      14.0%; Score 153; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.6e-23;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 GATAAGATAGGAGAGAGAGAGAGGCTGCGGCACCGAGGATCACCGTTGGGGGCTCCGC 223
Db 1 GATAAGATAGGAGAGAGAGAGGCTGCGGCACCGAGGATCACCGTTGGGGGCTCCGC 60

QY 224 GAGTACAGCAAAATAGTTTCTGAGAAAGTTGAAGCCAAAGAGAAACACATCAATGAG 283
Db 61 GAGTACAGCAAAATAGTTTCTGAGAAAGTTGAAGCCAAAGAGAAACACATCAATGAG 120

QY 284 GTTCGACAGCAAAATTTATTTCAGAGCTGAAGTCC 316
Db 121 GTTCGACAGCAAAATTTATTTCAGAGCTGAAGTCC 153

```

Search completed: February 6, 2005, 14:35:10  
Job time : 5151 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:23:21 ; Search time 697 Seconds  
(without alignments)

9249.065 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1084.2	99.6	1089	4	Aaf80144
2	442.2	40.6	885	12	ADQ36842
3	218	20.0	548	10	ADK71795
4	218	20.0	548	12	ADQ31159
5	168	15.4	168	4	Aaf80149
6	168	15.4	168	4	Aaf80148
7	157.6	14.5	1443	12	ADQ32852
8	155.2	14.3	1267	12	ADQ3084
9	155.2	14.3	1267	12	ADQ32144
10	153.6	14.1	1114	6	AAS96287
11	153.6	14.1	1189	6	AAS96319
12	153.6	14.1	1274	10	ADH59559
13	153.6	14.1	1274	10	ACC45100
14	153.4	14.1	1153	12	ADQ3626
15	153	14.0	153	4	Aaf80147
16	151.2	13.9	156	4	Aaf80146
17	150.8	13.8	1158	12	ADQ3082
18	150.8	13.8	1158	12	ADQ62142
19	150.8	13.8	1618	3	AAC39795
20	150	13.8	150	4	Aaf80150

21 138.2 12.7 900 3 AAS9702 DNA encod  
22 138.2 12.7 1245 12 ADO63363  
23 136.6 12.5 1193 2 AAZ34580  
24 136.2 12.5 1140 12 ADQ36854  
25 132.2 12.1 1041 12 ADQ36844  
26 127.8 11.7 962 12 ADO63216  
27 123 11.3 1317 12 ADO62851  
28 118.6 10.9 1131 6 AAS96304  
29 115 10.6 980 12 ADO63362  
30 104.6 9.6 425 6 AAS96415  
31 102.8 9.4 642 13 ADR5325  
32 99.2 9.1 2968 11 ACN88986  
33 98.8 9.1 2320 11 ADI32130  
34 98.8 9.1 4359 12 ADN05729  
35 97 8.9 1266 6 ABL65865  
36 97 8.9 1396 5 AAS92244  
37 97 8.9 2783 13 ACN38014  
38 97 8.9 2968 5 ABV26838  
39 97 8.9 2968 5 ABV20992  
40 95.2 8.7 665 2 AAZ34577  
41 93.6 8.6 870 6 AAS96410  
42 93.6 8.6 1442 6 AAS96332  
43 88 8.1 1832 2 AAV72862  
44 88 8.1 2203 4 ABL08071  
45 85 7.8 358 3 AAC56435

#### ALIGNMENTS

RESULT 1  
AAF80144  
ID AAF80144 standard; DNA; 1089 BP.

AC AAF80144;

DT 11-JUN-2001 (first entry)

DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.

XX E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
KW ss.

OS Triticum monococcum.

XX Key Location/Qualifiers

FT CDS 20..805

FT /\*tag= /a

FT /product= "E2F-dimerisation partner (DP) protein"

XX WO200121644-A2.

PD 29-MAR-2001.

PF 25-SEP-2000; 2000WO-EP009325.

PR 24-SEP-1999; 99ES-00002127.

PR 11-NOV-1999; 99ES-00002474.

XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

PI Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/28.

DR P-PSDB; AAB67762.

XX New isolated, enriched, cell free and/or recombinant nucleic acid useful  
PT for e.g. altering cell proliferation characteristic such as to alter  
PT plant cell, organ or tissue size.

PS Claim 13; Fig 1; 77pp; English.

CC The present sequence encodes a E2F-dimerisation partner (DP) protein. The  
CC protein acts as a plant E2F transcription factor. E2F and DP are two  
CC proteins that hetero-dimerise to form an active transcription factor that  
CC regulates G1 to S phase of the cell cycle, and later, the expression of  
CC genes required for S-phase progression. E2F and retinoblastoma protein  
CC also interact as a hetero-dimer in cells to suppress certain genes. This  
CC repression involves binding of the retinoblastoma protein to the E2F-DP  
CC dimer that is in turn bound to sites on DNA through the E2F DNA binding  
CC domain. DP proteins can be modulated to alter plant cell, organ or tissue  
CC shape, and particularly to alter cell proliferation characteristic such  
CC as to alter plant cell, organ or tissue size  
XX  
SQ Sequence 1089 BP; 367 A; 215 C; 261 G; 246 T; 0 U; 0 Other;

Query Match 99.6%; Score 1084.2; DB 4; Length 1089;  
Best Local Similarity 99.7%; Pred. No. 5.2e-207;  
Matches 1086; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAATTCGGCAGGAGCCGCGCTCCCGCGGCGGAGCTGCTCGCGCGCTACCGC 60  
Db 1 GAATTCGGCAGGAGCCGCGCTCCCGCGGCGGAGCTGCTCGCGCGCTACCGC 60  
Qy 61 CGCACTGGACCTGACCGGCGTGACATCTCGAAGCTTCCAGTGTCCCGCGCTCCCGA 120  
Db 61 CGCACTGGACCTGACCGGCGTGACATCTCGAAGCTTCCAGTGTCCCGCGCTCCCGA 120  
Qy 121 AGCGGCGGTAATCGGTCCTCAAGGAAAGGGGCTGTTGACCCCGGATAAAGATAGGAAGAA 180  
Db 121 AGCGGCGGTAATCGGTCCTCAAGGAAAGGGGCTGTTGACCCCGGATAAAGATAGGAAGAA 180  
Qy 181 GGAGAAGGCTCGCGACCGAGGATCACGGTGTGGGGCTCGCGAGTACAGCAAAATAGT 240  
Db 181 GGAGAAGGCTCGCGACCGAGGATCACGGTGTGGGGCTCGCGAGTACAGCAAAATAGT 240  
Qy 241 TTGTGAGAAAGTTGAAGCAAGGAAAGCAATCAATGAGGTTCGACACGAAATTTA 300  
Db 241 TTGTGAGAAAGTTGAAGCAAGGAAAGCAATCAATGAGGTTCGACACGAAATTTA 300  
Qy 301 TTCAGAGCTGAAGTCCATGGCACAATATGGTCAAGGGTTGATGAGAAGAAATATAGGCG 360  
Db 301 TTCAGAGCTGAAGTCCATGGCACAATATGGTCAAGGGTTGATGAGAAGAAATATAGGCG 360  
Qy 361 GAGAGTGTATGATGCTTTCACGCTTCATTCGCACTTCGTTGTTTGGAAAGAAAGAAA 420  
Db 361 GAGAGTGTATGATGCTTTCACGCTTCATTCGCACTTCGTTGTTTGGAAAGAAAGAAA 420  
Qy 421 GGAGATACGGTGGATGGGCTTTCAAATTCAGATATGAAAAATAAAGAGCTTGAGGA 480  
Db 421 GGAGATACGGTGGATGGGCTTTCAAATTCAGATATGAAAAATAAAGAGCTTGAGGA 480  
Qy 481 AGTTTCGTAAGAACTCGTCAACAGATAGGAACAAGAGGCACTCTCCAGGAAATCGA 540  
Db 481 AGTTTCGTAAGAACTCGTCAACAGATAGGAACAAGAGGCACTCTCCAGGAAATCGA 540  
Qy 541 AAAACAGTTTCATGATCTCCAAAACATCAAGTACGTAACCAACATCGGAAGCTCAGC 600  
Db 541 AAAACAGTTTCATGATCTCCAAAACATCAAGTACGTAACCAACATCGGAAGCTCAGC 600  
Qy 601 AGAGAAATGTAATGGCATCCGCTTCCATTCGTATGTGTCAGACATCTAGGAAGCAAG 660  
Db 601 AGAGAAATGTAATGGCATCCGCTTCCATTCGTATGTGTCAGACATCTAGGAAGCAAG 660  
Qy 661 GGTGGAATTCAGATTCAGATGACTCGAAGTTTGGCCCATTCAGTTCAATGTGTGACCC 720  
Db 661 GGTGGAATTCAGATTCAGATGACTCGAAGTTTGGCCCATTCAGTTCAATGTGTGACCC 720  
Qy 721 ATTCACATTCAGATTCATCTCAATCTTGGGGGTAAAGGCTAACACATAGGAAG 780  
Db 721 ATTCACATTCAGATTCATCTCAATCTTGGGGGTAAAGGCTAACACATAGGAAG 780  
Qy 781 AGCTGGCGCGCCACCTTCACTAGAGACTCAAGAAATATTACAAATGAATTAAGTGTT 840  
Db 781 AGCTGGCGCGCCACCTTCACTAGAGACTCAAGAAATATTACAAATGAATTAAGTGTT 840

Qy 841 AGAACTGGCAGCAGCCGAGTCTTTTCAGACAGCTATGTATAGCTATATATCTCATGAAA 900  
Db 841 AGAACTGGCAGCAGCCGAGTCTTTTCAGACAGCTATGTATAGCTATATATCTCATGAAA 900  
Qy 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAGATTTTAACCTGCAAAATTTG 960  
Db 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAGATTTTAACCTGCAAAATTTG 960  
Qy 961 TCTCCTTTTGTGCTAGCAGGTTTATAGTCTCAGATAGATTCATATATGTGCTGC 1020  
Db 961 TCTCCTTTTGTGCTAGCAGGTTTATAGTCTCAGATAGATTCATATATGTGCTGC 1020  
Qy 1021 TATGAAAACATTCATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1080  
Db 1021 TATGAAAACATTCATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1080  
Qy 1081 AAAAAAAGAA 1089  
Db 1081 AAAAAAAGAA 1089  
RESULT 2  
ADQ36842  
ID ADQ36842 standard; DNA; 885 BP.  
XX ADQ36842;  
AC AC  
XX AC  
DT 07-OCT-2004 (first entry)  
XX  
DE Cell proliferation-related nucleic acid sequence #1.  
XX cell proliferation related polypeptide; cell proliferation; senescence;  
KW differentiation; stress response; db.  
XX Oryza sativa.  
XX  
XX WO2004061122-A2.  
PN  
XX 22-JUL-2004.  
PD  
XX 23-DEC-2003; 2003WO-US041200.  
PF  
XX 26-DEC-2002; 2002US-0436565P.  
PR  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA  
XX Cooper B;  
XX WPI; 2004-534388/51.  
DR  
XX New nucleic acid molecule encoding a cell proliferation-related polypeptide, useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops.  
XX Claim 3; SEQ ID NO 1; 408pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule encoding a cell proliferation-related polypeptide. The nucleic acid molecule and the encoded polypeptide, and methods are useful for modulating cell proliferation, senescence, differentiation, development, and stress response in plants, and for producing enhanced food crops. The present sequence represents a cell proliferation-related nucleic acid sequence. The present sequence is published separately from the main body of the specification as EPO data.  
XX  
SQ Sequence 885 BP; 249 A; 203 C; 243 G; 190 T; 0 U; 0 Other;

Query Match 40.6%; Score 442.2; DB 12; Length 885;  
Best Local Similarity 79.8%; Pred. No. 8.7e-79;  
Matches 548; Conservative 0; Mismatches 133; Indels 6; Gaps 2;





AC ADO63159;  
XX  
XX  
DT 15-JUN-2004 (first entry)  
XX  
XX Transcription factor G2981 orthologous sequence, SEQ ID 1626.  
XX  
XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
XX osmotic stress tolerance; cold tolerance; heat tolerance;  
XX low nitrogen tolerance; low phosphate tolerance; fungal disease;  
XX glyphosate resistance; flowering; fertility; seed development; ds.  
XX  
XX Oryza sativa.  
XX  
XX WO2004031349-A2.  
XX  
XX 15-APR-2004.  
XX  
XX 18-SEP-2003; 2003WO-US030292.  
XX  
XX 18-SEP-2002; 2002US-0411837P.  
XX 17-DEC-2002; 2002US-0434166P.  
XX 24-APR-2003; 2003US-0465809P.  
XX  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX  
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;  
XX Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;  
XX WPI; 2004-330163/30.  
XX  
XX New recombinant polynucleotide encoding transcription factor  
XX polypeptides, useful for producing transgenic plants with advantageous  
XX properties compared to a reference plant.  
XX  
XX Claim 1; SEQ ID NO 1626; 510pp; English.  
XX  
XX The present invention relates to novel plant transcription factor  
XX proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The  
XX sequences can be used to produce transgenic plants, which overexpress  
XX (II), where the transgenic plant has an altered trait as compared to a  
XX non-transgenic plant or wild-type plant. The transgenic plant comprises  
XX an altered trait selected from increased tolerance to abiotic stress,  
XX increased tolerance to osmotic stress, increased tolerance to cold,  
XX increased germination in cold, increased tolerance to heat, increased  
XX germination in heat, increased tolerance to freezing conditions,  
XX increased tolerance to low nitrogen conditions, increased tolerance to  
XX low phosphate conditions, increased tolerance to disease, including  
XX fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
XX increased tolerance to multiple fungal pathogens, increased resistance to  
XX glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
XX increased sensitivity to ACC, altered sugar sensing, increased tolerance  
XX to sugars, altered carbon/nitrogen sensing, early flowering, late  
XX flowering, altered flower structure, loss of flower determinacy, reduced  
XX fertility, altered shoot meristem development, altered branching pattern,  
XX altered stem morphology, altered vascular tissue structure, reduced  
XX apical dominance, altered trichome density, altered trichome development,  
XX altered trichome structure, altered root development, altered shade  
XX avoidance, altered seed development, altered seed ripening, altered seed  
XX germination, slow growth, fast growth, altered cell differentiation,  
XX altered cell proliferation, altered cell expansion, altered phase change,  
XX altered senescence, abnormal embryo development, altered programmed cell  
XX death, lethality when overexpressed, altered necrosis patterns, dark  
XX plant size, increased biomass, large seedlings, dwarfed plants, dark  
XX green leaves, change in leaf shape, increased leaf size and mass, light  
XX green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,  
XX altered seed coloration, altered seed size, altered seed shape, large  
XX seed, increased leaf wax, increased leaf fatty acids, altered seed oil  
XX content, increased seed protein content, altered seedprenyl content,  
XX altered leaf prenyl lipid content, increased anthocyanin levels, and  
XX decreased anthocyanin levels. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 548 BP; 164 A; 118 C; 112 G; 154 T; 0 U; 0 Other;  
Query Match 20.0%; Score 218; DB 12; Length 548;  
Best Local Similarity 71.1%; Pred. No. 4.8e-34;  
Matches 391; Conservative 0; Mismatches 140; Indels 19; Gaps 7;  
QY 554 GATCTCCAAACATCAAGTTACCTGTAACCAACACTGGAAGCTCAGCAGAGATGTTAAT 613  
DB 1 GACCTTCCAGAAATATACATTACCAACAGGCTAGTCAGAGGCCA-CAGAAAGTGTAAAT 59  
QY 614 GGCATCCGCCTTCC-ATTTCGTATTGGTCAAGACATCTAGGAAAGCAGGCTGGAATTTGA 672  
DB 60 GGCATCTCTCTCCGGTTCCTATTGATCAAGACATCCGAAAGCAGGCTGGAATTTGA 119  
QY 673 GATTTCAGATGACTCGAAGTTTCCCATTTTCGAGTTCAATGGTGCACCATTCACATTGCA 732  
DB 120 GATTTCCGAAGATTCAAAGTTTCACGGTTTCGACTTCAACGGTGCACCATTCACCATGCA 179  
QY 733 TGATGATCTCTCAATCTCTTGGAGGGGTAAAGCGTAAACAGCATAGGAGAGCTGCCCGCG 792  
DB 180 TGATGATGATCAATCTCTTGAAGCCATCAGGCGGTAAACAAAGGAGAGCTGCCCTCTC 239  
QY 793 CACCTTCTCACT-AGAGACTCAAGAAATATTACAAATGAAATTTAAAGTG----TTAGAACTG 847  
DB 240 CATTCACTTAAAGAGGCAAGAAATCTGACAAACCATTTGAAGTGTCAAACTCAAACTG 299  
QY 848 GCACAGCGGATTCCTTTTGCACAGCTATGATATAGCTATA-----TATCCTCATGAAAA 900  
DB 300 GCACACCCAGTTTCTTTTGCACAGTTATGATATAGCTATAGCTTACCTTTCAGTATGAAA 359  
QY 901 CTGTGACTAGTTTATAGGACAGTCTCTCAGGCTTGAAGCA-TTTTAACTGCAAAATTTT 959  
DB 360 CTCGACTAGTTTATAGGACAGTCTCTCAGGCTTTTTCAGCATATAGCTTGCAGAAATTT 419  
QY 960 GTCTCTCTTTTGTGCTTAGCAGGTTATTAGTCTCAGATAGATGATTATATATGCTGCTG 1019  
DB 420 GCCTCTTTTGTGCTTAGCAGGTTA---GTCTGGAATAGTTTCTTTCGCTGTGCTG 475  
QY 1020 CTATGAAACATTTGATAGCAAAAAA-----TATCCTCATGAAAA 1079  
DB 476 TTATGATATAAATCAATTCGCGTTTCAATTGCAAGCTTGTGTAAGCTCCATTGATGAAA 535  
QY 1080 AAAAAA 1089  
DB 536 AAAAAA 545  
RESULT 5  
AAF80149  
ID AAF80149 standard; DNA; 168 BP.  
XX  
XX AAF80149;  
XX  
XX 11-JUN-2001 (first entry)  
XX  
XX DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.  
XX  
XX E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
XX S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
XX ss.  
XX  
XX Triticum monococcum.  
XX  
XX WO200121644-A2.  
XX  
XX 29-MAR-2001.  
XX  
XX 25-SEP-2000; 2000WO-EP009325.  
XX  
XX 24-SEP-1999; 99ES-00002127.  
XX 11-NOV-1999; 99ES-00002474.  
XX

PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX Gutierrez-Armenta C, Ramirez-Parra E;  
 PI WPI; 2001-257972/26.  
 XX P-PSDB; AAB67766.  
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful  
 DR for e.g. altering cell proliferation characteristic such as to alter  
 DR plant cell, organ or tissue size.  
 XX Disclosure; Page 73-74; 77pp; English.  
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation  
 XX partner (DP) protein. The protein acts as a plant E2F transcription  
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an  
 CC active transcription factor that regulates G1 to S phase of the cell  
 CC cycle, and later, the expression of genes required for S-phase  
 CC progression. E2F and retinoblastoma protein also interact as a hetero-  
 CC dimer in cells to suppress certain genes. This repression involves  
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn  
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and particularly  
 CC to alter cell proliferation characteristic such as to alter plant cell,  
 CC organ or tissue size  
 XX Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;  
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 485 CGTAAAGAACTCGTCAACAAGATTAGGACAAAGAGGAGGCTCTCCAGGAAATCGAAAA 544  
 DB 1 CGTAAGAACTCGTCAACAAGATTAGGACAAAGAGGAGGCTCTCCAGGAAATCGAAAA 60  
 QY 545 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604  
 DB 61 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 120  
 QY 605 AATGTTAATGGATCGCGCTTCATTCGTTATGTTGTTCAAGACATCTAGG 652  
 DB 121 AATGTTAATGGATCGCGCTTCATTCGTTATGTTGTTCAAGACATCTAGG 168  
 RESULT 6  
 AAF80148  
 ID AAF80148 standard; DNA; 168 BP.  
 XX AAF80148;  
 XX 11-JUN-2001 (first entry)  
 DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.  
 DE E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
 KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
 KW ss.  
 XX Triticum monococcum.  
 OS WO200121644-A2.  
 XX 29-MAR-2001.  
 PD 25-SEP-2000; 2000WO-EP009325.  
 XX 24-SEP-1999; 99ES-00002127.  
 PR 11-NOV-1999; 99ES-00002474.  
 XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PA Gutierrez-Armenta C, Ramirez-Parra E;  
 PI WPI; 2001-257972/26.  
 XX P-PSDB; AAB67766.  
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful  
 DR for e.g. altering cell proliferation characteristic such as to alter  
 DR plant cell, organ or tissue size.  
 XX Disclosure; Page 74-75; 77pp; English.  
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation  
 XX partner (DP) protein. The protein acts as a plant E2F transcription  
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an  
 CC active transcription factor that regulates G1 to S phase of the cell  
 CC cycle, and later, the expression of genes required for S-phase  
 CC progression. E2F and retinoblastoma protein also interact as a hetero-  
 CC dimer in cells to suppress certain genes. This repression involves  
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn  
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and particularly  
 CC to alter cell proliferation characteristic such as to alter plant cell,  
 CC organ or tissue size  
 XX Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;  
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 485 CGTAAAGAACTCGTCAACAAGATTAGGACAAAGAGGAGGCTCTCCAGGAAATCGAAAA 544  
 DB 1 CGTAAGAACTCGTCAACAAGATTAGGACAAAGAGGAGGCTCTCCAGGAAATCGAAAA 60  
 QY 545 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604  
 DB 61 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 120  
 QY 605 AATGTTAATGGATCGCGCTTCATTCGTTATGTTGTTCAAGACATCTAGG 652  
 DB 121 AATGTTAATGGATCGCGCTTCATTCGTTATGTTGTTCAAGACATCTAGG 168  
 RESULT 6  
 AAF80148  
 ID AAF80148 standard; DNA; 168 BP.  
 XX AAF80148;  
 XX 11-JUN-2001 (first entry)  
 DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.  
 DE E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
 KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
 KW ss.  
 XX Triticum monococcum.  
 OS WO200121644-A2.  
 XX 29-MAR-2001.  
 PD 25-SEP-2000; 2000WO-EP009325.  
 XX 24-SEP-1999; 99ES-00002127.  
 PR 11-NOV-1999; 99ES-00002474.  
 XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PA Gutierrez-Armenta C, Ramirez-Parra E;  
 PI WPI; 2001-257972/26.  
 XX P-PSDB; AAB67766.  
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful  
 DR for e.g. altering cell proliferation characteristic such as to alter  
 DR plant cell, organ or tissue size.  
 XX Disclosure; Page 73-74; 77pp; English.  
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation  
 XX partner (DP) protein. The protein acts as a plant E2F transcription  
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an  
 CC active transcription factor that regulates G1 to S phase of the cell  
 CC cycle, and later, the expression of genes required for S-phase  
 CC progression. E2F and retinoblastoma protein also interact as a hetero-  
 CC dimer in cells to suppress certain genes. This repression involves  
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn  
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and particularly  
 CC to alter cell proliferation characteristic such as to alter plant cell,  
 CC organ or tissue size  
 XX Sequence 168 BP; 59 A; 19 C; 44 G; 46 T; 0 U; 0 Other;  
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 317 ATGCGACATATTGGTCAAGGGTTTGATGAGAAGAATATTAGCGGAGAGTGTATGATGCT 376  
 DB 1 ATGCGACATATTGGTCAAGGGTTTGATGAGAAGAATATTAGCGGAGAGTGTATGATGCT 60  
 QY 377 TTCAACGGTCTCATTTGACACTTCGTGTTATTGCAAAAGAAAAAGAGAGATACGGTGGATG 436  
 DB 61 TTCAACGGTCTCATTTGACACTTCGTGTTATTGCAAAAGAAAAAGAGAGATACGGTGGATG 120  
 QY 437 GGCTTTTCAAAATTCACAGATATGAAATAAAGAAAGCTTCAGGAAGTT 484  
 DB 121 GGCTTTTCAAAATTCACAGATATGAAATAAAGAAAGCTTCAGGAAGTT 168  
 RESULT 7  
 ADO62852  
 ID ADO62852 standard; DNA; 1443 BP.  
 XX ADO62852;  
 XX 15-JUL-2004 (first entry)  
 DE Transcription factor G2981/2982 orthologous sequence, SEQ ID 1319.  
 DE Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
 KW osmotic stress tolerance; cold tolerance; heat tolerance;  
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;  
 KW glyphosate resistance; flowering; fertility; seed development; ds.  
 XX Glycine max.  
 OS WO2004031349-A2.  
 XX 15-APR-2004.  
 PD 18-SEP-2003; 2003WO-US030292.  
 XX 18-SEP-2002; 2002US-0411837P.  
 PR 17-DEC-2002; 2002US-0434166P.  
 PR 24-APR-2003; 2003US-0465809P.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;  
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;

XX WPI; 2004-330163/30.

XX

PT New recombinant polynucleotide encoding transcription factor

PT polypeptides, useful for producing transgenic plants with advantageous

PT properties compared to a reference plant.

XX

PS Claim 1; SEQ ID NO 1319; 510pp; English.

XX

CC The present invention relates to novel plant transcription factor

CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The

CC sequences can be used to produce transgenic plants, which overexpress

CC non-transgenic plant or wild-type plant. The transgenic plant comprises

CC an altered trait selected from increased tolerance to abiotic stress,

CC increased tolerance to osmotic stress, increased tolerance to cold,

CC increased germination in cold, increased tolerance to heat, increased

CC germination in heat, increased tolerance to freezing conditions,

CC increased tolerance to low nitrogen conditions, increased tolerance to

CC low phosphate conditions, increased tolerance to disease, including

CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,

CC increased tolerance to multiple fungal pathogens, increased resistance to

CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,

CC increased sensitivity to ACC, altered sugar sensing, increased tolerance

CC to sugars, altered carbon/nitrogen sensing, early flowering, late

CC flowering, altered flower structure, loss of flower determinacy, reduced

CC fertility, altered shoot meristem development, altered branching pattern,

CC altered stem morphology, altered vascular tissue structure, reduced

CC apical dominance, altered trichome density, altered trichome development,

CC altered trichome structure, altered root development, altered shade

CC avoidance, altered seed development, altered seed ripening, altered seed

CC germination, slow growth, fast growth, altered cell differentiation,

CC altered cell proliferation, altered cell expansion, altered phase change,

CC altered senescence, abnormal embryo development, altered programmed cell

CC death, lethality when overexpressed, altered necrosis patterns, increased

CC plant size, increased biomass, large seedlings, dwarfed plants, dark

CC green leaves, change in leaf shape, increased leaf size and mass, light

CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,

CC altered seed coloration, altered seed size, altered seed shape, large

CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil

CC content, altered seed protein content, altered seedprenyl content,

CC altered leaf prenyl lipid content, increased anthocyanin levels, and

CC decreased anthocyanin levels. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 1443 BP; 419 A; 273 C; 343 G; 403 T; 0 U; 5 Other;

Query Match 14.5%; Score 157.6; DB 12; Length 1443;

Best Local Similarity 56.1%; Pred. No. 6.8e-22;

Matches 348; Conservative 0; Mismatches 254; Indels 18; Gaps 2;

QY 154 TGTGACCGGATTAAGATAGGAAGAGGCTGCGCACCGAGGATCACCGGTG 213

DB 257 TGTGGCAAGAGAAAAGAGAGCGGTCAACGGCTGTGGACCTGATAGAGTGAAG 316

QY 214 GGGCTCCGCGAGTACAGCAAAATAGTTTGTGAAAAGTTGAAGCCAAAGAGAAACAAC 273

DB 317 AGGTCTCCGCCAATTTAGTATGAAGTGTGTGAGAAGGTAGAAAGCAGGGGAAGAACAC 376

QY 274 ATACAATGAGTTTCACAGCAAAATTTATTCAGACT-----GAAGTCCAT 318

DB 377 ATACAATGAGTGGCGGATGAGCTTTGTGTGAATTTCTGAACCAAGCAATAGTGAAT 436

QY 319 GGACATATTTGTTCAAGGTTTGTATGAGAAATATTAGCGGAGAGTGTATGATGCTTT 378

DB 437 GCGCCCTGATCAGCAACATATGATGAAAAAATCATCGCGAAGGCTCTATGATGCTCT 496

QY 379 CAAGCTTCTATTGCACCTTCGTGTTATTGCAAAAAGAAAAAGAGATACGGTGGATGGG 438

DB 497 GAAGCTTCTATGCGCAATGGATATATTTCGAAGGACAAAAAGAAATTCATATGGAGGG 556

QY 439 CTTTCAAATTACAGATATGAAAAAATAAGAGCTTGGAGAGTTTCGTAAAGAACTCGT 498

DB 557 CTTCTCTGTAATCTGTAATGTAATTAAGAGCAAGAGCGGCTTGGGCTCAG 616

QY 499 CAAACAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGATGATCT 558

DB 617 GAATGAATTTGAAAAGAAAAACAGCCCTATCTGCAAGGAGCTTGAGGAGCAATTCGTAGTCT 676

QY 559 CAAAACATCAAGTTAGGTAACCAAAACACTGGAAGCTCAGCAGAGATGTTTAAAT---GG 615

DB 677 TCAGAACCTTATTCAACGAAATGAGCAGTTATATAGCTCAGGAAATCTCCCAAGTGAGG 736

QY 616 CATCCGCTTCCATTGCTTCAAGACATCTAGGAAAGCAAGGTTGAAATTTGAGAT 675

DB 737 TGTATCTTTACCTTTATTTTGTACAGACACTCTCTCATGCAACTGTGGAAGTGAAT 796

QY 676 TTCAAGATGACTCGAAGTTTGCCCATTTCAATGGTGCACCAATTCATTTGCATGA 735

DB 797 ATCAGAAGATATGCAAGTTTGTTCATTTGATTTCAATAGCACTCTTTTGAGCTGCATGA 856

QY 736 TGATCTCTCAATCTTGAGG 755

DB 857 CGACAATTATGTTCTCAAGG 876

## RESULT 8

ADO63084

ID ADO63084 standard; DNA; 1267 BP.

XX ADO63084;

XX 15-JUL-2004 (first entry)

XX Transcription factor G2982 coding sequence, SEQ ID 1551.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;

KW osmotic stress tolerance; cold tolerance; heat tolerance;

KW low nitrogen tolerance; low phosphate tolerance; fungal disease;

KW glyphosate resistance; flowering; fertility; seed development; ds.

XX Arabidopsis thaliana.

XX WO2004031349-A2.

XX 15-APR-2004.

XX 18-SEP-2003; 2003WO-US030292.

XX 18-SEP-2002; 2002US-0411837P.

XX 17-DEC-2002; 2002US-0434166P.

XX 24-APR-2003; 2003US-0465809P.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;

PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;

XX WPI; 2004-330163/30.

DR P-PSDB; ADO63085.

PT New recombinant polynucleotide encoding transcription factor

PT polypeptides, useful for producing transgenic plants with advantageous

PT properties compared to a reference plant.

XX Claim 1; SEQ ID NO 1551; 510pp; English.

XX The present invention relates to novel plant transcription factor

CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The

CC sequences can be used to produce transgenic plants, which overexpress

CC (II), where the transgenic plant has an altered trait as compared to a

CC non-transgenic plant or wild-type plant. The transgenic plant comprises

CC an altered trait selected from increased tolerance to abiotic stress,

CC increased tolerance to osmotic stress, increased tolerance to cold,

CC increased germination in cold, increased tolerance to heat, increased  
 CC germination in heat, increased tolerance to freezing conditions,  
 CC increased tolerance to low nitrogen conditions, increased tolerance to  
 CC low phosphate conditions, increased tolerance to disease, including  
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
 CC increased tolerance to multiple fungal pathogens, increased resistance to  
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late  
 CC flowering, altered flower structure, loss of flower determinacy, reduced  
 CC fertility, altered shoot meristem development, altered branching pattern,  
 CC altered stem morphology, altered vascular tissue structure, reduced  
 CC apical dominance, altered trichome density, altered trichome development,  
 CC altered trichome structure, altered root development, altered shade  
 CC avoidance, altered seed development, altered seed ripening, altered seed  
 CC germination, slow growth, fast growth, altered cell differentiation,  
 CC altered cell proliferation, altered cell expansion, altered phase change,  
 CC altered senescence, abnormal embryo development, altered programmed cell  
 CC death, lethality when overexpressed, altered necrosis patterns, increased  
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark  
 CC green leaves, change in leaf shape, increased leaf size and mass, light  
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,  
 CC altered seed coloration, altered seed size, altered seed shape, large  
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil  
 CC content, altered seed protein content, altered seedprenyl content,  
 CC altered leaf prenol lipid content, increased anthocyanin levels, and  
 CC decreased anthocyanin levels. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1267 BP; 420 A; 225 C; 267 G; 355 T; 0 U; 0 Other;

Query Match 14.3%; Score 155.2; DB 12; Length 1267;  
 Best Local Similarity 55.7%; Pred. No. 2e-21;  
 Matches 351; Conservative 0; Mismatches 258; Indels 21; Gaps 2;

QY 131 AATCGCGTCCAAAGGAGGGGGCTGTGACCGGATAAAGATAGGAAGGAGGCT 190  
 DB 239 ACTCCAGTGAGAGGAAATGATGTGTGATGATGATCTGAAATGGTCAGAGAGAA 298  
 QY 191 GCGGACCGAGGATCAACCGGTTGGGGCTCCGCGAGTACAGCAAAATAGTTGTGAGAA 250  
 DB 299 GGGCAATCAAGAACTTCGAGGCGGGCTTCGTCAATTCACTGTTATGTTTGTGAGAG 358  
 QY 251 GTTGAGCCAAAGGAGCAACATACATAGCTGTCAGCAAAATTTATTCAGAGCTG 310  
 DB 359 TTGGAGCCAAAGGAGATTAATCTACTTCAAGGAGGTTGCGACGAAATTTATTCAGATTT 418  
 QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGTTTGTATGAGAGAT 352  
 DB 419 GCCACAATTAAGCAAAAGCAGAGAGCCCTTGAATGAATAAGTACATGAGAGAGAAC 478  
 QY 353 ATTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTTGCACTTCGTGTTATTCAGAAA 412  
 DB 479 ATAAGCGGAGAGTCTACGATGCGTCAATGTGTTCAATGCGTGTGATTTATTCAGAG 538  
 QY 413 GAAAAAAGAGATACGGTGGATGGGCTTTCAAAATACAGATATGAAAAAATAAAGAG 472  
 DB 539 GATAAAAAGGAATCCGGTGGAAAGACTTCTTATTACCTGCAAAAAGGATGTGGAAGAA 598  
 QY 473 CTTGAGGAAGTGTGAAGAACTCGTCAACAGATTAGGACAGAGCACTCTCCAG 532  
 DB 599 GTCAAGATGATGTAATAAGTTATGAGAGTGTGCAAAAAGAGCTGCTTTCTTAA 658  
 QY 533 GAAATCGAAAAACAGTTTGATGATCTCCAAAACATCAAGTTAGCTTAACAA---ACACTG 589  
 DB 659 GAGTTGAGAGAAAGTCTCAAGTCTTGAGAGCTTTATGTGAGAAATCAAGAGTGGT 718  
 QY 590 GAAAGCTCAGCAGAGAAATTTATGATGATCCCGCTTCATTCGTTATTTGTCAGACATCT 649  
 DB 719 GTGAAGACTCAAGGCCCCAGAGAGGATTTACCTTACCATTCTTACTTGTGAGACAAAC 778

QY 650 AGGAAGCAGAGGTGGAAATTCAGATTTCAGATGACTCGAAGTTTGCCATTTTCGAGTTC 709  
 DB 779 CCTCAGCAGTAGTCGAATTCGAGATTCGAGATTCGAGATTCGAGATTCGAGATTC 838  
 QY 710 AATGGTGACCAATTCACATTCGATGATGATGAT 739  
 DB 839 AATAGCACACCTTTCTCGGTCCATGATGAT 868

# RESULT 9

AD062144  
 ID AD062144 standard; DNA; 1267 BP.

XX AC AD062144;

DT 15-JUL-2004 (first entry)

Transcription factor G2981 coding sequence, SEQ ID 611.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
 KW osmotic stress tolerance; cold tolerance; heat tolerance;  
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;  
 KW glyphosate resistance; flowering; fertility; seed development; ds.

XX Arabidopsis thaliana.

XX PN WO2004031349-A2.

XX PD 15-APR-2004.

XX PF 18-SEP-2003; 2003WO-US030292.

XX PR 18-SEP-2002; 2002US-0411837P.

XX PR 17-DEC-2002; 2002US-0434166P.

XX PR 24-APR-2003; 2003US-0465809P.

(MEND-) MENDEL BIOTECHNOLOGY INC.

XX PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;  
 PI Riechmann JU, Haake V, Dubell AN, Keddle JS, Sherman BK;

XX WPI; 2004-330163/30.

XX P-PSDB; AD062145.

XX New recombinant polynucleotide encoding transcription factor  
 PT polypeptides, useful for producing transgenic plants with advantageous  
 PT properties compared to a reference plant.

XX Claim 1; SEQ ID NO 611; 510pp; English.

XX The present invention relates to novel plant transcription factor  
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The  
 CC sequences can be used to produce transgenic plants, which overexpress  
 CC (II), where the transgenic plant has an altered trait as compared to a  
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises  
 CC an altered trait selected from increased tolerance to abiotic stress,  
 CC increased tolerance to osmotic stress, increased tolerance to cold,  
 CC increased germination in cold, increased tolerance to heat, increased  
 CC germination in heat, increased tolerance to freezing conditions,  
 CC increased tolerance to low nitrogen conditions, increased tolerance to  
 CC low phosphate conditions, increased tolerance to disease, including  
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
 CC increased tolerance to multiple fungal pathogens, increased resistance to  
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance  
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late  
 CC flowering, altered flower structure, loss of flower determinacy, reduced  
 CC fertility, altered shoot meristem development, altered branching pattern,  
 CC altered stem morphology, altered vascular tissue structure, reduced  
 CC apical dominance, altered trichome density, altered trichome development,  
 CC altered trichome structure, altered root development, altered shade  
 CC avoidance, altered seed development, altered seed ripening, altered seed  
 CC germination, slow growth, fast growth, altered cell differentiation,





QY 191 GCAGCACCAGGATCACCGTGTGGGGCTCCGCGAGTACAGCAAAATAGTTGTGAAA 250  
 Db 146 GGGCAATCAAGACTTCTGGAGCGGGCTTGGTCAATTCAGTGTATGGTTTGCAGAG 205  
 QY 251 GTTGAAGCCAAAGAGAACACATACATCAATAGAGTGTGCAGACGAAATTTATTCAGAGCTG 310  
 Db 206 TTGGAAGCCAAAGAGATAACTACTTACAGAGGTTGCAGACGAAATTTATTCAGATTTT 265  
 QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGGTTTGTATGAGAGAAT 352  
 Db 266 GCCACATTAAGCAAAAGCAGAGAGCCCTTGAATGAAATGAGTACAATGAGAAGAAC 325  
 QY 353 ATTAGCGGAGAGTGTATGCTTTCAACGTTCTCAATTCGACTTCGTGTATTGCAAAA 412  
 Db 326 ATAAGCGGAGAGTCTACGATCGCTCAATGTGTTCATGGGTTGGATATTATTCGAAG 385  
 QY 413 GAAAAAAGAGATACGCTGGATGGGCTTTCAAAATACAGATATGAAAAAATAAGAG 472  
 Db 386 GATAAAAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAAGGATGTGAAGAA 445  
 QY 473 CTTGAGGAGTTCGTAAGAACTCGTCAACAGATTAGGACAGAGGCACTCTCCAG 532  
 Db 446 GTCAGATGATCGTATAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAA 505  
 QY 533 GAAATCGAAAAACAGTTTGATGATCTCCAAAACATCAAGTTACGTAACCAA---ACACTG 589  
 Db 506 GAGTTCAGAGAAAGGCTCAAGTCTTGAGAGCTTATGTGCGAGAAATCAAGAGATGGTT 565  
 QY 590 GAAAGCTCAGCAGAGAAATTAATGGAATCCGCTTCCATTCGTATGGTCAAGACATCT 649  
 Db 566 GTGAAGACTCAAGGCCAGCAGAGAGATTTACCTTACCATTCATCTTCTGAGACAAAC 625  
 QY 650 AGGAAGCAAGGCTGGAATGAGATTCAGATGACTCGAAGTTTGCCCATTTGCGATTC 709  
 Db 626 CTTACGCGATAGTCGAAATCGAGATTTCTGAAGATGATGCAACTGTACACCTCGACTTC 685  
 QY 710 AATGGTGCACCATTCACATTCGATGAT 739  
 Db 686 AATGACACACCTTCTCGGTCCATGAT 715

RESULT 11  
 AAS96319  
 ID AAS96319 standard; cDNA; 1189 BP.  
 XX  
 AC AAS96319;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE Arabidopsis cDNA encoding cell cycle protein CCP16.  
 XX  
 KW Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;  
 KW plant growth regulator; plant development; abiotic stress; biotic stress;  
 KW nutrient deprivation; pathogen attack; crop yield.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W0200185946-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 14-MAY-2001; 2001WO-1B001307.  
 XX  
 PR 12-MAY-2000; 2000US-0204045P.  
 XX  
 PA (CROP-) CROPDESIGN NV.  
 XX  
 PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;  
 XX  
 DR WPI; 2002-062249/08.  
 DR P-PSDB; AAU72529.  
 XX

PT New cell cycle protein and nucleic acid molecule encoding it useful for  
 PT regulating cell cycle progression in plants and for identifying  
 PT modulators which are useful as herbicides or plant growth regulators.  
 XX  
 PS Claim 38; Fig 16; 316pp; English.  
 XX  
 CC The invention relates to a novel cell cycle protein (CCP) and the  
 CC polynucleotides encoding them. CCP is useful for identifying a compound  
 CC which modulates the activity of the polypeptide and which binds to the  
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
 CC maize, tomato, alfalfa, oilseed rape, soybean, thaliana, rice, wheat,  
 CC nucleic acid and polypeptide molecules are useful as modulating agents in  
 CC regulating cell cycle progression in plants. CCP is useful to treat  
 CC disorders characterised by insufficient or excessive production of CCP  
 CC protein or production of CCP protein forms which have decreased or  
 CC aberrant activity. Compounds that bind to or modulate the activity of CCP  
 CC polypeptide are useful as herbicides or plant growth regulators. The  
 CC polynucleotide is useful for modifying cell fate, plant development,  
 CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S/G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, nodule function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due to  
 CC environmental conditions, including abiotic stress such as cold, nutrient  
 CC deprivation, heat, drought, salt stress, or biotic stress such as  
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate  
 CC plant architecture, plant quality traits, plant reproduction and seed  
 CC development, endoreduplication in storage cells, storage tissues and/or  
 CC storage organs of plants or its parts. CCP is useful as an immunogen to  
 CC generate antibodies. CCP protein is useful to screen for naturally  
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP  
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to  
 CC modulate CCP activity. The present sequence encodes a CCP protein of the  
 CC invention  
 XX  
 SQ Sequence 1189 BP; 400 A; 202 C; 260 G; 327 T; 0 U; 0 Other;

Query Match 14.1%; Score 153.6; DB 6; Length 1189;  
 Best Local Similarity 55.6%; Pred. No. 4.1e-21;  
 Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;  
 QY 131 AATCGGTCGCAAGGAGGGGCTGTGACCCGGAATAAGATAGGAAGAGGAGGCT 190  
 Db 161 ACTCCAGTGAGAGGAAATTTGATTGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 220  
 QY 191 GCGGCACCGAGATCACCGTGTGGGGCTCCGCGAGTACAGCAAAATAGTTGTGAGAA 250  
 Db 221 GGCAATCAAGAACTTCTGGAGCGGGCTTCGTAATTCAGTGTATGGTTTGCAGAG 280  
 QY 251 GTTGAAGCCAAAGAGAACACATACATCAATAGAGTGTGCAGACGAAATTTATTCAGAGCTG 310  
 Db 281 TTGGAAGCCAAAGAGATTAATCTACTTACAGGAGGTTGCAGACGAAATTTATTCAGATTTT 340  
 QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGGTTTGTATGAGAAGAAT 352  
 Db 341 GCCACATTAAGCAAAAGCAGAGAGCCCTTTCATGAAATGAGTACATGAGAAGAAC 400  
 QY 353 ATTAGCGGAGAGTGTATGCTTTCAACGTTCTCATTCGACTTCGTGTATTGCAAAA 412  
 Db 401 ATAAGCGGAGAGTCTACGATCGCTCAATGTGTTCATGGCGTTGGATATTATTCGAAG 460  
 QY 413 GAAAAAAGAGATACGCTGGATGGGCTTTTCAAAATTTACAGATATGAAAAAATAAGAG 472  
 Db 461 GATAAAAGGAAATCCGGTGGAAAGAGCTTCTTATTCCTGCAAAAGAGATGTGGAAGAA 520  
 QY 473 CTTGAGGAAAGTTGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCTCCAG 532  
 Db 521 GTCAAGATGGATCGTAAATAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTCTTAA 580

QY 533 GAAATCGAATAAAGAGTTTGAATGATCTCCAAAACATCAAGTTACGTAAACAA---ACACTG 589  
Db 581 GAGTTGAGAGAAAGGTCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 640  
QY 590 GAAAGCTCAGCAGAGAAATGTTAATGCGATCCGCTTCCATTGCTATGTTGTCAGACATCT 649  
Db 641 GTGAAGACTCAAGGCCAGAGAGGATTTACCTTACCATTCTTCTTACCTGAGACAAAC 700  
QY 650 AGGAAAGCAAGGTGGAATTTGAGATTTTCCAGATGACTCGAAGTTTGCCTTCCATTTCGAGTTTC 709  
Db 701 CCTCAGCAGTAGTCGAATCGAGATTTCTGAGATATGCAACTTGTACACCTCGACTTC 760  
QY 710 AATGGTGACCAATTCATTCGATGATGAT 739  
Db 761 AATAGCACACCTTTCTCGGTCCATGATGAT 790

RESULT 12  
ADH59559  
ID ADH59559 standard; DNA; 1274 BP.  
XX AC  
XX ADH59559;  
XX DT 25-MAR-2004 (first entry)  
XX DE Dimerisation partner encoding sequence.  
XX DE E2F transcription factor; Gene therapy; stress resistance;  
XX KW differentiation signal; ds.  
XX KW Arabidopsis thaliana.  
XX OS  
XX WO2003025185-A1.  
XX PN  
XX PD 27-MAR-2003.  
XX PF 12-SEP-2002; 2002WO-EP010236.  
XX PR 14-SEP-2001; 2001EP-00870198.  
XX PA (CROP-) CROPDESIGN NV.  
XX PI Beekman T, De Veylder L, Inze D, Mironov V, Broekaert W;  
PI Dillen W, Frankard V;  
XX DR WPI; 2003-371815/35.  
XX DR P-PSDB; ADH59560.  
XX PT Modifying cell number, architecture and yield of plants comprises  
PT modulating the expression and/or the activity of members of the plant E2F  
PT transcription factor family.  
XX PS Disclosure; SEQ ID NO 3; 41pp; English.  
XX CC The present invention relates to increasing the cell number of specific  
CC cell types, specific tissues or specific organs in a plant comprises  
CC modulating the expression and/or the activity in the specific cell types,  
CC specific tissues or specific organs of the plant E2F transcription  
CC factor. The E2F transcription factor or its homologue, derivative or  
CC fragment, is useful for prolonging the period of cell division in certain  
CC cells and tissues, for increasing the size of cotyledons, for enhancing  
CC cell proliferation after seed germination, for enhancing stress  
CC resistance of seedlings, for obtaining seedlings with enhanced vigor, for  
CC obtaining plants that have increased growth or number of organs, for  
CC obtaining plants having more cells in a particular tissue, for obtaining  
CC plants having an increased size of organs or having increased yield, for  
CC stimulating differentiated cells to re-enter the cell cycle, for  
CC overriding the cell differentiation signals, or for altering cell shape.  
CC The present sequence represents the dimerisation partner encoding  
XX sequence.

Sequence 1274 BP; 420 A; 228 C; 271 G; 355 T; 0 U; 0 Other;

Query Match 14.1%; Score 153.6; DB 10; Length 1274;  
Best Local Similarity 55.6%; Pred No. 4.2e-21;  
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;  
QY 131 AATGCGTCCAAAGGAAGGGGCTGTTGACCCGGATAAGATAGGAAGAAGAGAGGCT 190  
Db 246 ACTCCAGTGAGAGGAAATTTGATTTGATGATGATCTGAAATTTGGATTCAGAGAAGAA 305  
QY 191 GGGGACCGAGATCACCGGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAA 250  
Db 306 GGGCAATCAAGAACTTCTGGAGGGGGCTTCGTCAAATCAGGTGTTATGTTTGTGAGAAG 365  
QY 251 GTTGAAGCCAAAGGAAGAACATACATGATGAGTTGCGAGACGAAATTTATTTCAGAGCTG 310  
Db 366 TTGGAAGCCAAAGAGTAATCTTCAAGGAGTTGCGAGACGAAATTTATTTCAGATTTT 425  
QY 311 AAGTCCAT-----GGCACATATTGTTCAAGGGTTTGCATGAGAAGAT 352  
Db 426 GCCACAAATTAAGCAAAACGACAGAAAGCCTTTGGAATGAAATGAGTACATGAGAAGAAC 485  
QY 353 ATTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTTGCATCTTCGTGTTATGCAAAA 412  
Db 486 ATAAGCGGAGAGTCTACGATCGCTCAATGTTGTTCAATGCGCTTGGATTTATTTGCAAGG 545  
QY 413 GAAAAAAGGAGATACGGTGGATGGGCTTTCAAAATTTACAGATATGAAAAAATAAAGAAAG 472  
Db 546 GATAAAGGAATTCGGTGGAAAGGACTTCTTATACCTGCAAAAAGGATGTGGAAGAA 605  
QY 473 CTTGAGGAAGTTCTGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAG 532  
Db 606 GTCAGATGATCGTAAATAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTCTTTAA 665  
QY 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTAGCTAACCAA---ACACTG 589  
Db 666 GAGTTGAGAGAAAGGCTTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGTT 725  
QY 590 GAAAGCTCAGCAGAGAAATGTTAATGGCATCCGCTTCCATTTCGTATTTGGTCAAGACATCT 649  
Db 726 GTGAGACTCAAGGCCAGCAGAGGATTTACCTTACCATTCTTCTTACCTGAGACAAAC 785  
QY 650 AGGAAAGCAAGGTGGAATTTGAGATTTTCAAGTATTCAGATGACTCGAAGTTTCCCAATTCGAGTTTC 709  
Db 786 CCTCAGCAGTAGTCGAAATTCGAGATTTCTGAGATATGCACTTGTACACCTCGACTTC 845  
QY 710 AATGGTGACCAATTCATTCGATGATGAT 739  
Db 846 AATAGCACACCTTTCTCGGTCCATGAT 875

RESULT 13  
ACC45100  
ID ACC45100 standard; cDNA; 1274 BP.  
XX AC  
XX ACC45100;  
XX DT 10-JUN-2003 (first entry)  
XX DE Arabidopsis thaliana DP encoding cDNA SEQ ID NO:3.  
XX KW Arabidopsis thaliana; modulation; endoreduplication; plant;  
XX KW transgenic plant; cell cycle; E2F; DP; gene; ss.  
XX OS Arabidopsis thaliana.  
XX FH Key Location/Qualifiers  
XX CDS 174..1052  
XX FT /\*tag= a  
XX FT /product= "DP protein"  
XX PN WO2003018818-A2.  
XX PD 06-MAR-2003.





CC	altered cell proliferation, altered cell expansion, altered phase change,	
CC	altered senescence, abnormal embryo development, altered programmed cell	
CC	death, lethality when overexpressed, altered necrosis patterns, increased	
CC	plant size, increased biomass, large seedlings, dwarfed plants, dark	
CC	green leaves, change in leaf shape, increased leaf size and mass, light	
CC	green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,	
CC	altered seed coloration, altered seed size, altered seed shape, large	
CC	seed, increased leaf wax, increased leaf fatty acids, altered seed oil	
CC	content, altered seed protein content, altered seedprenyl content,	
CC	altered leaf prenyl lipid content, increased anthocyanin levels, and	
CC	decreased anthocyanin levels. Note: The sequence data for this patent did	
CC	not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 1153 BP; 337 A; 225 C; 267 G; 324 T; 0 U; 0 Other;	
	Query Match 14.1%; Score 153.4; DB 12; Length 1153;	
	Best Local Similarity 55.1%; Pred. No. 4.5e-21;	
	Matches 354; Conservative 0; Mismatches 271; Indels 18; Gaps 2;	
QY	162 CGGATTAAGATAGGAAGAGGAGGCTCGGCACCGAGGATCACCGGTTGGGGCTCC 221	
DB	169 CTGGGAAGAAAGAGCTCAGCGTCTACTGGAGGGATAGAGTGTAGAGGACTCA 228	
QY	222 GCGAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATG 281	
DB	229 GACAGTTTAGCATGAAAGTTGTGAGAAAGTGGAAAGCAAGGAAGAACTACGTATAATG 288	
QY	282 AGTTGACAGCAAAATTTATTCAGAGCTGAAGTCCATATTTGGTCAAGG---- 336	
DB	289 AGTTGACAGTGAAGTGTAGTCTAGTCTGAGTTTCTGATGCTCAATAGTTGGAGGCTCAG 348	
QY	337 -----GTTTGATGAGAAGAAATATTAGCGCGAGAGTGTATGATGCTTCAAGCTTC 386	
DB	349 ATCGAACAATATGATGAGAAGAACTACAGCAGCAGTCTACGATGCTCTGACGTAC 408	
QY	387 TCATGTGCACTTCGTTTATTTCGCAAAAGAAAGGAAGATACGTTGATGGGCTTTCAA 446	
DB	409 TTATGGCTATGGATATCATTTCTAAGATATAAAGGAATACAGTGGAAAGGATTACCAC 468	
QY	447 ATTACAGATATGAAAATAAAGAGCTTGAAGGAAGTTCTGAAAGAACTCGTCAACAAGA 506	
DB	469 GCACGTATGCAATATGATGAGGAGCTTAAGACTGAGCGTCTTAACCTGAGAATAGGA 528	
QY	507 TTAGGAACAAGAGGCACTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACA 566	
DB	529 TTGAAAAGAAAGCAGCTTATTTAGAGAACTTGAAGATCAATATGTAGGGCTTCAAAACC 588	
QY	567 TCAAGTTAGTTAACCAAAACACTGGNAAGCTCAGCAGAGAAATGTTA---ATGGCATCGCC 623	
DB	589 TCATAAAAACGCAATGATCAGTTGTATGCTCAGGCAATGCTCTAGTGGTGGTGGCTT 648	
QY	624 TTCCATTTCGTTATGCTCAAGACATCTAGGAAGCAAGGGTGGAAATTCAGATTTTCAGATG 683	
DB	649 TACCGTTTATTTAGTGAGACTCGTCTCATGCTACAGTGAAGTGAATATACAGAAG 708	
QY	684 ACTCGAAGTTGCCATTTTCGAGTTCAATGGTGCACCAATTCACATTCGATGATCTCT 743	
DB	709 ATATGAGCTGTGTGATTTTCGACTTTCAACAGCAGCACTCCGTTTGGAGCTGATGACGATAAT 768	
QY	744 CAATCTTTCAGGGGTGAGGCGTAAACAGCATAGGAAGAGCTGG 786	
DB	769 ATATCTCTCAAGCAATGAATTTTTCGGGAAGATCAATGACGG 811	
RESULT 15		
AAF80147		
ID	AAF80147 standard; DNA; 153 BP.	
XX		
AC	AAF80147;	
XX		
DT	11-JUN-2001 (first entry)	

XX	DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.	
DE	E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;	
XX	S phase; cell cycle; retinoblastoma protein; alter cell proliferation;	
KW	ss.	
KW		
XX	Triticum monococcum.	
OS		
XX	WO200121644-A2.	
PN		
PD	29-MAR-2001.	
XX		
XX	25-SEP-2000; 2000WO-EP009325.	
PF		
XX	24-SEP-1999; 99ES-00002127.	
PR		
XX	11-NOV-1999; 99ES-00002474.	
PR		
XX	(CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.	
PA		
XX	Gutierrez-Armenta C, Ramirez-Parra E;	
PI		
XX	WPI; 2001-257972/26.	
DR	P-PSDB; AAB67765.	
DR		
XX	New isolated, enriched, cell free and/or recombinant nucleic acid useful	
PT	for e.g. altering cell proliferation characteristic such as to alter	
PT	plant cell, organ or tissue size.	
PT		
XX	Disclosure; Page 72-73; 77pp; English.	
PS		
XX	AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation	
CC	partner (DP) protein. The protein acts as a plant E2F transcription	
CC	factor. E2F and DP are two proteins that hetero-dimerise to form an	
CC	active transcription factor that regulates G1 to S phase of the cell	
CC	cycle, and later, the expression of genes required for S-phase	
CC	progression. E2F and retinoblastoma protein also interact as a hetero-	
CC	dimer in cells to suppress certain genes. This repression involves	
CC	binding of the retinoblastoma protein to the E2F-DP dimer that is in turn	
CC	bound to sites on DNA through the E2F DNA binding domain. DP proteins can	
CC	be modulated to alter plant cell, organ or tissue shape, and particularly	
CC	to alter cell proliferation characteristic such as to alter plant cell,	
CC	organ or tissue size	
XX		
SQ	Sequence 153 BP; 55 A; 25 C; 47 G; 26 T; 0 U; 0 Other;	
	Query Match 14.0%; Score 153; DB 4; Length 153;	
	Best Local Similarity 100.0%; Pred. No. 3.6e-21;	
	Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	164 GATAAAGATAGGAAGAGGAGGAGGCTGCGGCACCGAGGATCACCGTTGGGGGCTCCGC 223	
DB	1 GATAAAGATAGGAAGAGGAGGAGGCTGCGGCACCGAGGATCACCGTTGGGGGCTCCGC 60	
QY	224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAGCCCAAGGAGAACACATCAATGAG 283	
DB	61 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAGCCCAAGGAGAACACATCAATGAG 120	
QY	284 GTTGCAGACCAAAATTTATTTCAGAGCTCAAGTCC 316	
DB	121 GTTGCAGACCAAAATTTATTTCAGAGCTCAAGTCC 153	
Search completed: February 6, 2005, 13:05:23		
Job time : 702 secs		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:47:57 ; Search time 4249 Seconds  
(without alignments)  
9755.705 Million cell updates/sec

Title: US-10-088-830-1  
Perfect score: 1089  
Sequence: 1 gaattggcagcgagcgca.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	614.4	56.4	649	6	CD882927 F1.111L02
2	421.6	38.7	765	6	CA764996 AF53-Rpf
3	394.2	36.2	454	6	CD892664 GI18.121L7
4	376.4	34.6	607	2	BE456002 HVM5MG001
5	373.8	34.3	612	7	CR289243 CR289243
6	368.2	33.8	1119	9	CL959517 OsIFCC003
7	286.8	26.3	665	6	CA078334 SCLAM100
8	277	25.4	548	2	BE511883 946054A10
9	264.2	24.3	634	7	CF484481 POL1_25C
10	237.8	21.8	702	7	CR286882 CR286882
11	218	20.0	548	4	BI802607 H084D10 E
12	160.6	14.7	836	9	CL909692 OA_Aba000
13	152.8	14.0	992	7	CK278661 EST724739
14	150.8	13.8	1403	3	EX831191 Arabidops
15	148.2	13.6	1418	3	ENS09YMC
16	143	13.1	1568	3	CNS09ZIR
17	142	13.0	680	6	CA290160 SCARFL801
18	138.2	12.7	1245	3	AY108383 Zea mays
19	135	12.4	501	5	BE680618 BX680618
20	135	12.4	696	9	CL618454 OR_Bba001
21	135	12.4	751	9	CL773898 OR_Bba008
22	135	12.4	789	9	CL773941 OR_Bba008
23	135	12.4	867	9	CL838498 OR_Cba006
24	131.2	12.0	743	7	CN215980 29828 Sus

25	127.8	11.7	498	1	A1939068
26	124.6	11.4	583	5	BU579005
27	123	11.3	643	6	C96290
28	122.2	11.2	627	1	AU237921
29	121.8	11.2	550	4	BG040894 NXSI_116
30	121	11.1	870	6	CD573622 UCRPT01_0
31	119.8	11.0	783	1	AJ804378
32	119	10.9	577	2	BE020396 em43901.Y
33	118.4	10.9	728	4	BG873645 MEST8-E05
34	118.2	10.9	852	7	CO116984 GR_EB019
35	116	10.7	673	7	CK751888 eca01-7cs
36	115.6	10.6	760	4	BM406319 EST580646
37	115	10.6	980	4	BG320803 Zm04_0190
38	113.2	10.4	706	7	CF482425 POL1_6 H0
39	112.8	10.4	645	6	CB917390 VVD02067
40	112.8	10.4	646	6	CB919100 VVD041C12
41	111.4	10.2	790	8	AQ577328 nxb0090F
42	111	10.2	454	5	BU012590 QGJ2F12.Y
43	111	10.2	552	7	CK754643 pam01-7ms
44	110.6	10.2	640	6	CB919899 VVD056H09
45	110	10.1	733	6	CD813391 BN15.019H

#### ALIGNMENTS

RESULT 1  
CD882927  
LOCUS F1.111L02F010430 F1 Triticum aestivum cDNA clone F111L02, mRNA  
DEFINITION sequence.  
ACCESSION CD882927  
VERSION CD882927.1 GI:32645392  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Genoplante.  
TITLE Genoplante, a major partnership french program in plant genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Genoplante  
Genoplante  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).

#### FEATURES

source  
Location/Qualifiers  
1..649  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="recital"  
/db\_xref="taxon:4565"  
/clone="F111L02"  
/tissue\_type="leaf one"  
/clone\_lib="F1"

#### ORIGIN

Query Match 56.4%; Score 614.4; DB 6; Length 649;  
Best Local Similarity 96.8%; Pred. No. 1.8e-123;  
Matches 627; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 290 GACGAATTTATTTCAGAGCTGAAGTCCATGGGCACATATTGGTCAAGGTTTGATGAGAAG 349  
DB 2 GATGAATTTATTTCAGAGCTGAAGTCCATGGGCACATATTGGTCAAGGTTTGATGAGAAG 61  
QY 350 AATATTAGCGGAGAGTGATGATGCTTTCAAGCTTCTCATTCGACTTCGTGTTATGCA 409  
DB 62 AATATTAGCGGAGAGTGATGATGCTTTCAAGCTTCTCATTCGACTTCGTGTTATGCA 121



```
RESULT 3
CD892664
LOCUS
DEFINITION
G118.121J23P010725 G118 Triticum aestivum cDNA clone G118121J23,
mRNA sequence.
ACCESSION
CD892664
VERSION
CD892664.1 GI:32663198
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 454)
Genoplante, a major partnership french program in plant genomics
UNPUBLISHED (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
Location/Qualifiers
1. .454
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G118121J23"
/tissue_type="grain (118 degrees per day after
pollination)"
/clone_lib="G118"

ORIGIN
source
Query Match 36.2%; Score 394.2; DB 6; Length 454;
Best Local Similarity 96.7%; Pred. No. 2e-75;
Matches 413; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 608 GTTAATGGCATCGCCTTCATTTCGTTATGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 667
Db 11 GGTATGGCATCGCCTTCATTTCGTTATGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 70

QY 668 ATTGAGATTCAGATCACTCGAGTTTGCCATTTCGAGTTCAATGGTGCACCATTCACA 727
Db 71 ATTGAGATTCAGATCACTCGAGTTTGCCATTTCGAGTTCAATGGTGCACCATTCACA 130

QY 728 TTGCATGATGATCTCTCAATCTCTGAGGGGTAAAGCGGTAAACAGCATAGAGAGCTGGC 787
Db 131 TTGCATGATGATCTCTCAATCTCTGAGGGGTAAAGCGGTAAACAGCATAGAGAGCTGGC 190

QY 788 CGGCGCACCCCTTCACATAGACATCAAGATATATTAACAATGAATATAAAG-TGTTAGAACT 846
Db 191 CGGCGCACCCCTTCACATAGACATCAAGATATATTAACAATGAATATAAAGTTGTGCAACT 250

QY 847 GCACAGCCGGATCTTTTGCACAGCTATGTATAGCTATATATCTTCATGAAAACCTTGAC 906
Db 251 GCACAGCCGGATCTTTTGCACAGCTATGTATAGCTATATATCTTCATGAAAACCTTGAC 310

QY 907 CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGAGATTTTAACTGCAAAATTTTGTCTCT 966
Db 311 CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGAGATTTTAACTGCAAAATTTTGTCTCT 370

QY 967 TTTTGTGCTAGCAGCTTATAGGCTCAGATAGATGATTCATATATGCTGCTATGAA 1026
Db 371 TTTTGTGCTAGCAGCTTATAGGCTCAGATAGATGATTCATATATGCTGCTATGAA 430

QY 1027 AACATTG 1033
Db 431 AACATTG 437

ORIGIN
Query Match 34.6%; Score 376.4; DB 2; Length 607;
BE456002 607 bp mRNA linear EST 22-OCT-2001
HVSMEG0019D17f Hordeum vulgare pre-anthesis spike EST library
HVCNDA0008 (white to yellow anther) Hordeum vulgare subsp. vulgare
cDNA clone HVSMEG0019D17f, mRNA sequence.
BE456002
BE456002.3 GI:16318905
EST
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 607)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
Yu,X., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
Fenton,R.D., Close,S.J., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex pre-anthesis spike cDNA library
UNPUBLISHED (2001)
On Jul 26, 2000 this sequence version replaced gi:13155022.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 141
Seq primer: AATTACCTCCTCACTAAGG
High quality sequence stop: 585.
Location/Qualifiers
1. .607
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEG0019D17f"
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCNDA0008 (white to yellow anther)"
/notes="Vector: lambdaZAP; Site1: EcoRI; Site2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spike with awns trimmed were collected at white, green and
yellow anther stages (Fenton). Total RNA was prepared from
each pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagmids. These steps were performed in the TJ Close lab
(Choi) at the University of California, Riverside.
Phagmids were plated and picked at the Clemson University
Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins
and Wing) Plasmid DNA preparations, DNA sequencing and
sequence analysis were performed at CUGI (Wing, Yu,
Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence
has been trimmed to remove vector sequence and contains a
minimum of 100 bases of phred value 20 or above. For more
details on library preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
```

```
Best Local Similarity 87.9%; Pred. No. 1.5e-71;
Matches 457; Conservative 0; Mismatches 56; Indels 7; Gaps 4;

QY 9 CACGAGCCGCAATGGCGCTCCCGCGGGGAGCTGCTGGCGCGCTACCGCGCACTGG 68
Db 28 CTGACACCGCAATGGCGCTCCCGCGGAGGTGCTGCTGGCGCGTTCGCGCGCACTGG 87
QY 69 ACTGACCGCGGTGCACATCTCGAAGCTTCAGAGTTCCTCCCGCGCTTCGCGGAA---GCGG 125
Db 88 ACTGACCGCGGTGCACATCTCGAAGCTTCAGAGTTCCTCCCGGTTCGCGGTTCCCAACGCGGG 147
QY 126 GCGGTATCGGTCCTAAGGAGGCGGCTGTTGACCGGATTAAGATAGGAAGGAGA 185
Db 148 GCGGCAATCGGTCCTAAGGAGGAGGCGGCTTACCGCGGATTAAGATAGGAAGGAGA 207
QY 186 AGCTCGGCGCACGAGGATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTG 245
Db 208 AGCTCGGCGCACATAGATCACCGGTTGGGGGCTCCGCGAATACAGCAAAATAGTTTGTG 267
QY 246 AGAAAGTTGAAGCAAGGAAGAACACATACATCAATGAGGTTCAGACGAAATTTATTCAG 305
Db 268 AGAAAGTTGAAGCAAGGAAGAACACATACATCAATGAGGTTCAGGATGCGGATTAATTCAG 327
QY 306 AGCTGAAGTCCATGGGACATATTTGTCAGGGTTGATGAGAAGATATTTAGCGCGAGAG 365
Db 328 AGCTGAAGTCCATGGGACATATTTGTCAGGGTTGATGAGAAGATATTTAGCGCGGAG 387
QY 366 TGATGATGCTTTCAAGCTTCTCATTCGACCTTCGTTATTTGCAAAAGGAAAAAGGAGA 425
Db 388 GGTATGATGCTTTCAAGCTTCTCATTCGACCTTCATGATATCGCAAAAGAACAAAGGATA 447
QY 426 TAGCGTGGATGGG-CTTTCAAAATACAGATATGAAAAA--ATAAGAGCTTCGAGGAG 482
Db 448 ATCGCGGATGGGCGCTTTCAAAATACAGATATGAAAAAGCTTAAACAAAGTTGGAGGAG 507
QY 483 TT-CGTAAAGAACTCGTCAACAGATTAGCAACAGAGG 521
Db 508 TTCCCAAAAGAACTCGTCAACAGATTATGCAACAGGAG 547

RESULT 5
CR289243
LOCUS
DEFINITION
CR289243 Oryza sativa library (Han B) Oryza sativa cDNA clone
p704a10p5, mRNA sequence.
CR289243
VERSION
KEYWORDS
SOURCE
EST.
Oryza sativa
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 612)
Han,B., Feng,Q., Huang,Y.C., Ying,K., Li,Y., Guan,J.P., Zhu,J.J.,
Zhao,Q., Hu,X., Liu,Y.L., Wu,J., Yu,Z., Chen,L., Fan,D.L.,
Weng,Q.J., Zhang,L., Lu,X.Q., Yu,S.L., Liu,X.H., Lu,T.P.,
Zhang,Y.J., Lu,Y., Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X.,
Zhang,L., Lan,L.F., Chen,W., Wu,S.A. and Xue,Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
Location/Qualifiers
1. .612
/organism="Oryza sativa"
/mol_type="mRNA"

FEATURES
source
/db_xref="taxon:4530"
/clone="p704a10p5"
/clone_lib="Oryza sativa library (Han B)"

ORIGIN
Query Match 34.3%; Score 373.8; DB 7; Length 612;
Best Local Similarity 79.7%; Pred. No. 5.7e-71;
Matches 441; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 226 GTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAAACAACATCAATGAGGT 285
Db 1 GTTCAGCAAGATAGTTTCTAAGAAAGTTGAGCCAAAGAGAAACACATATAATGAGGT 60
QY 286 TCCAGACGAAATTTATTCAGAGCTCAAGTCCATGCGACATATTCGTCAGGGTTCATGATCA 345
Db 61 TCCGATTAAGATTTTTCGGGAGCTGAAGTCCATTAACGAGAACGGTCTGGAGTTGATGA 120
QY 346 GAAGAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTGTTAT 405
Db 121 GAAGAATATTAGCGGAGGTTATATGATGCTTTCAATGTCTCATTTGCAATTCGTGTTAT 180
QY 406 TCCAAAGAAAAAGGAGATACGGTGGATGGGCTTTTCAAAATTCAGATATGAAAAAT 465
Db 181 TCCAAAGAAAAAGGAGATAAAGTGGATGGGCTTTACTAATTTAGATACGAAAAAGAT 240
QY 466 AAAGAGCTTGGAGGAGTTTCGTAAGAGACTCGTCAACAGATTAGGAACAAGAGGCACT 525
Db 241 ACAGAGTTGGAGGAGTTTCACAAAGACTCATCAGGATCAAGATTAAGAGAGGT 300
QY 526 CTCCAGGAAATTCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAAC 585
Db 301 TCTCCAGGAAATTTGAAAGCAGTTTGTATGACCTTCAGAATATTACATTAACGCAACCG 360
QY 586 ACTGAAAGCTCAGCAGAGAAATGTTAATGGCATCGGCTTCATTCGTTATTCGTCAGAC 645
Db 361 TAGTCAGAGCGCAGCAGAAAGTGAATGGCATCTCTTCCTTCGCTATTATTGATCAAGAC 420
QY 646 ATCTAGGAAGCAAGGTCGAAATTTGAGATTTTCAGATGACTCGAAGTTTGGCCCATTTCCA 705
Db 421 ATCCGAAAGCAAGGTCGAAATTTGAAATTTGGAGAGTTTCCAACTTTGCACGGTTCCA 480
QY 706 GTTCAATGTGACCAATTCACATTCGATGATCTCTCAATCTTCCTTGGAGGGGTAAAGCG 765
Db 481 CTTCACTGTGCAACCAATCACCATGATGATATATCAATCCTTGAATATCATCATGCG 540
QY 766 TAACAGCATAGGA 778
Db 541 TACCACCAACCGA 553

RESULT 6
CR289517
LOCUS
DEFINITION
CR289517 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
CL959517
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 1119)
Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
```



Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: [chenchen@genomics.org.cn](mailto:chenchen@genomics.org.cn)  
Rice genomic sequence.  
Class: exon-trapped.

## FEATURES

```
1. 1119
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequence"
```

## ORIGIN

Query Match 33.8%; Score 368.2; DB 9; Length 1119;  
Best Local Similarity 78.8%; Pred. No. 1e-69;  
Matches 439; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

248	Qy	AAAGTTGAAGCCAAAGGAAGAAACAACATACAAATGAGGTTGCAGACGAAATTTATTCTAGAG	307
562	Db	AACGATGAAGTCTATACAAGCTCAGGAGAGCTTATAGTTGCCGATGAGATTTTTTTCGGAG	621
308	Qy	CTGAAGTCCATGGCACATATTTGGTCAAGGGTTTCATCGAGAAGAATAATTAGSCGGAGAGTG	367
622	Db	CTGAAGTCCATATACGCAAGACGGTCTCGAGTTTGATGAGAAGAATAATTAGSCGGAGGGTA	681
368	Qy	TATGATGCTTTTCAACGTTTCTCATTGCACTTCGTCGTTATTGCAAAAGAAAAAAGAGGATA	427
682	Db	TATGATGCTTTCAATGTGCTCATTTGCCAATTCGTCGTATTGCAAAAGATAAAAAAGGAGATA	741
428	Qy	CGGTGGATGGGCCCHTTCCAAATTACAGATATGAAAAAAATAAAGAAGCTTGTAGGAAAGTTCGT	487
742	Db	AAGTGGATGGGCCCHTTACTAATTATAGATATCGAAAAGGATACAGAAGTTTGGAGGAATTCAC	801
488	Qy	AAAGAACTCGTCAACAAGATTAGGAAACAAGAGGCACCTCTCCACGAGGAAATCGGAAAAACAG	547
802	Db	AAAGAACTCATACCCAGGATCAAGATTAAGAAGAAGCTTCTCCAGGAAATTTGAAAAGCAG	861
548	Qy	TTTGATGATCTCCAAAAACATCAAGTTTACGTAAACCAAAACACTGGAAAAGCTCAGCAGAGAAT	607
862	Db	TTTGATGACCTTCAGAAATATTACATTACGCAACCAACGCGTAGTCAGAGGCCAGCAGAAAGT	921
608	Qy	GTTAAATGGCATCCGCGCTTCCATTTCGTATTTGGTCAAGACATCTAGGAAGCAAGGGTGGAA	667
922	Db	GTTAAATGGCATCCCTCTCTCCGTTCTTATTGATCAAGACATCCCAGAAAGCAAGGGTGGAA	981
668	Qy	ATTGAGATTTTCAGATGACTCGAAGTTTGGCCATTTCGAGTTTCATGTTGTCACCAATTCACA	727
982	Db	ATTGAGATTTTCGGAAGATTCAAAGTTTTCGACGGTTTCGACTTCACACGGTGCACCAATTCACC	1041
728	Qy	TTGCATGATGATCTCTCAATCCTTTGAGGGGGTAAAGGGGTAAACAGCATAGGAAGAGCTGGC	787
1042	Db	ATGCATGATGATGATCAATCCTTTGAAGCCATCAGGGGTAAACAACAAGAGAGAGCTGGC	1101
788	Qy	CGGCGCCACCTTCTACTA	804
1102	Db	CTCTCCATTCACCCCTTA	1118

## RESULT 7

CA078334  
LOCUS  
DEFINITION  
665 bp mRNA linear  
SRLAM1006A02.g AM1 Saccharum officinarum cDNA clone  
5', mRNA sequence.  
EST 23-SEP-2003  
SRLAM1006A02

**ACCESSION**

VERSION CA078334.1 GI:3493060

## KEYWORDS

**SOURCE** Saccharum officinarum

**SOURCE**  
**ORGANISM**

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

complex.  
1 (bases 1 to 665)  
Vettore, A.B., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
<http://www.bcccenter.fcav.unesp.br>  
Plate: 006 row: A column: 02  
Seq primer: T7 Promoter Primer.

## FEATURES

## Source

```

1. 565
/organism="Saccharum officinarum"
/mol_type="mrna"
/db_xref="taxon:4547"
/clone="SCLRAM1006A02"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site 1: SalI; Site 2:
NotI. An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
Superscript Plasmid System kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CU-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucfest.lad.ic.unicamp.br/public
```

## ORIGIN

Query Match	26.3%	Score 286.8	DB 6	Length 665
Best Local Similarity	73.0%	Pred. No. 5.5e-52		
Matches 427	Conservative 0	Mismatches 142	Indels 16	Gaps 4
Qy	50	GC CGCTAC CGCCG CACTGGACCTGACCGGCGTGCACATCTCCGAAGCTTCCAGTGTCCCC	109	
Db	87	GAGCGGGCGCGCAGCCGAGCTCACAGCATCGGCATCATCGAAGAGCTGGCTCCCT	146	
Qy	110	CCGCTTCCCGAA-----GCGGCGGTAAATGCGGTCCAAAGGAAGGGGCTGTTCAC	160	
Db	147	CCGCTTCTTGACGGCGGAAGCGTCAGGGCAATGCGGCGAGAGNGAGAGGGTGAT	206	
Qy	161	CCGGATAAAGATAGGAAGAGGAAGGCTGCGGCACCGAGGATCACCGGTTGGGGCTC	220	
Db	207	ACGAAACAAGACGAGGAGGAAGGCTGGTATGCAGCGGATCGCTGGTGGGGCTC	266	
Qy	221	CGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCMAAGGAGAAACAATACAAT	280	
Db	267	CGCGAGTTGAGCAAGATAGTTTCTAAAAAAGTTTGAGACCMAAGGACGGACTACTTACAAT	326	
Qy	281	GAGGTTGCGACGAAATTTATTCAGAGCTGAAGTCCATGCCACATATTGCTCAAGGGTTT	340	
Db	327	GAGTTTGCAGATGAATTTGTGGGAGCTGAAG--CTGACTCTTATTTGTCGAAGTGT	383	
Qy	341	GATGAGAAGAAATATTAGCGGAGAGGTATGATGCTTTCAACGTTCTCATTTGCATTCGT	400	
Db	384	GATGAGAGNATATCAGGAGGAGGTATGATGCTTTTAACTGTCTAATCGCACTACGT	443	
Qy	401	GTTATTGCAAAAGAAAAAGGAGATACGGTGGATGGGCTTTCAAATTTACAGATGAA	460	
Db	444	GTTATTACAAAAGACCAAAAGGAGATAAAGTGATGGGCTTGCTTAATCCAAATGAA	503	
Qy	461	AAATTAAGAGCTTGAGGAAGTTCTGTAAAGAACTCGTCAACAGATTAGGAACAAGAAG	520	
Db	504	AAGATA---AAGTTGGAGAAATCCGCAAGAACTTATGATCAGGATAAGATAAGAAA	560	

```

QY 521 GCACTCTCCAGAAATCGAAACAGTTTGATGATCTCAAAACATCAAGTTACGTAAC 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 AAACCTCTCCAGAAATCGAACACAGTTTGATGACCTTCAGAAATATCAAGTTCCGCAAC 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 581 CAAACACTGGA-AAGCTCAGCAGAGAATGTTAATGCGATCCGCT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 CAGTTACTACGAAGCCAGCAAGAGTGCGAATGGATCAGCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
BE511883
LOCUS BE511883
DEFINITION 946064A10.y1 946 - tassels primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION BE511883
VERSION BE511883.1 GI:9733131
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 548)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946064 row: A column: 10.
FEATURES
    source
        1..548
            Location/Qualifiers
                1..548
                    /organism="Zea mays"
                    /mol_type="mRNA"
                    /cultivar="OH43"
                    /db_xref="taxon:4577"
                    /tissue_type="tassels"
                    /dev_stage="just after the transition from vegetative to
                    inflorescence development"
                    /lab_host="XLOLR"
                    /clone_lib="946 - tassels primordium prepared by Schmidt
                    lab"
                    /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
                    Site 2: XhoI; George Chuck dissected immature tassels
                    between 1mm and 3mm. Sharon Stanfield prepared the cDNA
                    library in HybriZAP. Sample insert size range was 350 bp
                    to 3 Kb with a 1 Kb average."

ORIGIN
    Query Match 25.4%; Score 277; DB 2; Length 548;
    Best Local Similarity 73.9%; Pred. No. 7.5e-50;
    Matches 397; Conservative 0; Mismatches 125; Indels 15; Gaps 3;

QY 53 GCTACCGCCGCTGACCTGACCGCGTGCACATTTCCGAAGCTTCCAGTGTCCCGCCG 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 GCTGCGCGCGCAGCCGAGCTCACCAGCTCGCATCAGCAAGAGCTGGATCCCTCCG 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 113 CTTCCCGAA-----CGCGCGGTATGCGGTCCAAAGAGGGGGTGTGACCCG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CTTCCCGAGCGCGCAAGCGTCGCGGCAATGCGCTGAGGAGAGAGAGGGGTGATAAG 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 164 GATAGATAGGAGAGGAGAGGCTGCGGCACCGAGGATCACCGTTGGGGCTCCGC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 AACAAAGCGGAGGAGGTAGGCTGGTACGAGCGGATCCCGGTGGGGGCTCCGC 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAAACATCAATGAG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 198 GAGTTCCAGCAAGATAGTTTCTAAAGGTTTGAGACCAAGGACGGACTACTTACAGTGAG 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 GTTGAGAGCAAAATTTATTCAGAGCTGAAGTCCATGGGCATATTTGGTCAAGGTTTGAT 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 GTTGAGAGTGAATTTGTGGAGAGCTGAAGT---TGACTCTTTATTGGTCAAGAGTTTGAT 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 GAGAAGATATTAGGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTCACACTTCCTGTT 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 GAGAAGATATCAGGAGGAGAGTGTATGATGCTTTTATGTGCTAATTCGACTACGTGTT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 ATTGCAAAAGAAAAAGGAGATACGGTGGATGGCGCTTTCAAATTCAGATATGAAAAA 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 ATCACAAGAGTAAAAAGAGATAAAATGATGGGCTTCTCTAAATTCCAATATGAAAAAG 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 ATAAGAAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCA 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 ATA--AAGTTGGAGGAAACTCGCAAGAACTCATGATCAGGATTAAAGAACAGAAAAA 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 CTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACATCAAGTTACGTAAC 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 CTTCTCCAGGAAATGAAAGACAGCTTGATGACCTCCAGATATCAAGTTTCGCAAC 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
CF484481
LOCUS CF484481
DEFINITION POLI_25_C08.g1_A002 Pollen Sorghum bicolor cDNA clone
POLI_25_C08_A002 5', mRNA sequence.
ACCESSION CF484481
VERSION CF484481.1 GI:34513350
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 634)
Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
Sun,F., Sullivan,R., Eastman,A. and Pratt,L.H.
An EST database from Sorghum: pollen
Unpublished (2003)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).
FEATURES
    source
        1..634
            Location/Qualifiers
                /organism="Sorghum bicolor"
                /mol_type="mRNA"
                /cultivar="BTx623"
                /db_xref="taxon:4558"
                /clone="POLI_25_C08_A002"
                /lab_host="DH10B-Ti phage-resistant E. coli"
                /clone_lib="Pollen"
                /note="Organ: Pollen; Vector: pME18S-FL3; Site 1: XhoI;
                Site 2: XhoI; The library was prepared from polyA+ RNA
                from pollen at the late vacuolated stage of
                development. Pollen was harvested from greenhouse-grown
                panicles of sorghum line Btx623. Panicles were removed
                from the flag leaf prior to emergence, when no detectable
                amylase is present in pollen of male-fertile lines. This

```



stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

Query Match	24.3%	Score	264.2	DB	7	Length	634
Best Local Similarity	70.1%	Pred. No.	4.7e-47				
Matches	385	Conservative	0	Mismatches	158	Indels	6
Gaps	2						
QY	15	CCGCAATGGCGCCTCCCGCGCGGAGTGTCTGCGGCGCTACCGCGCACTGGACCTGA	74				
Db	92	CCGTAATGGCGCCCTCTGCGCGACGCGGCGGCGGACGCGACTCACAGACACGGCGCA	151				
QY	75	CCGCGGTGCACATTCTCGAAGCTTCCAGTGTCCCCCGCTTCCCGAAGCGGCGGTAATG	134				
Db	152	TCACGCAAGGAGCGCGGCTCCCTCCGCTTCCTCTCGGACGCGGAAGGCTCGGCGGCAATG	211				
QY	135	CGGTCCAAAGGAGGGGCTGTTCACCGCGATAAGATAGGAAGAGGAGAGAGCTCGGG	194				
Db	212	CGGTTGAGGAGAGGAGGAGGGTGTATACGAACAAGAGCGGAGGAAGGAGAGCTGGTA	271				
QY	195	CACCGAGATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTGTGAGAAAGTTG	254				
Db	272	CGCAGCGGATCGCGGGTGGGGCTTCCGAGTTTCAGCAAGTATGTTCTAAAAAGTTG	331				
QY	255	AAGCCAAAGGAAGAACACATPACAATGAGGTTGCAGACGAAATTTATTCAGAGCTGAAGT	314				
Db	332	AGAGCAAGGACCGACTACTTACAATGAGGTTGCAGATGAAATTTGTGGGAGCTGAAG-	390				
QY	315	CCATGGCACATATTGTTCAAGGTTTCATGAGAGAAATATTAGCGGAGAGTGTATGATG	374				
Db	391	--CTGACTCTTATTGGTCAAGAGTTTATGAGAAAGAAATATCAGGAGAGAGTGTATGATG	448				
QY	375	CTTTCAAGTTCTCATTTGCATCTTCGTGTTATTGCAAAAGAAAAAAGAGATACGGTGA	434				
Db	449	CTTTTAATGTGCTAATCGCACTACGTGTTATTACAAAAGACCAAAAGGAGATAAAGTGA	508				
QY	435	TGGGCTTTCAAATTACAGATATGAAAAAATAAAGAGCTTGAAGAGAGTTTCGTAAAGAAC	494				
Db	509	TGGGCTTGTCTAACTTCCAAATATGAATACATA--AAAGTTGGAGGAACCTCGCAAGAAC	565				
QY	495	TCGTCAACAAGATTAGGAAACAAGAGGCACCTCCAGGAAATCGAAAAACAGTTTCATG	554				
Db	566	TCATGATTATGGTTAAGAACAGAAAAAACTTCTCCAGGAAATTTGAACGACAGTTTCATG	625				
QY	555	ATCTCCAAA	563				
Db	626	ACCTCCAGA	634				

RESULT 10				
CR286882				
LOCUS	702 bp	mRNA	linear	EST 27-FEB-2004
DEFINITION	CR286882	Oryza sativa library (Han B)	Oryza sativa cDNA clone	
			p730d01p5, mRNA sequence.	

Accession	CR280802	GI:44673448
VERSION	CR286882.1	
KEYWORDS	EST.	
SOURCE	<i>Oryza sativa</i>	
ORGANISM	<i>Oryza sativa</i>	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; <i>Oryza</i> .	
REFERENCE	1 (bases 1 to 702)	
AUTHORS	Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J., Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.	
TITLE	Rice cDNA EST clone	

**JOURNAL  
COMMENT**

Unpublished (2003)  
Contact: Han Bin  
National Center for Gene Research  
Chinese Academy of Sciences  
500# Cao Bao Road, Shanghai 200233, China  
Email: bhan@ncgr.ac.cn  
Clone requests: bhan@ncgr.ac.cn  
This is rice cdna est.clon  
Web site: <http://www.ncgr.ac.cn>.  
Location/Qualifiers  
1. 702  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/db\_xref="taxon:4530"  
/clone="p73d01p5"  
/clone.lib="Oryza sativa library"

<b>FEATURES</b>	<b>SOURCE</b>
-----------------	---------------

```

1. .02
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="p730d01p5"
/clone_lib="Oryza sativa library (Han B)"

```

## ORIGIN

Query Match	21.8%	Score 237.8	DB 7	Length 702
Best Local Similarity	75.7%	Pred. No. 2,7e-41		
Matches 373	Conservative 0	Mismatches 112	Indels 8	Gaps 6
Qy	192	CGGCACCGAGGATCACCGGTTGGGGGCTCGCGAGTACAGCAAAATAGTTTGTGACGAAG	251	
Db	26	CGGGCCAGAGGATCCCGGTTGGGGGCTCGTGAGTTCAGCAAGATAGTTTCTAAGAAAG	85	
Qy	252	TTGAAGCCAAAGGAAGAAACAAATACAATGAGGTTGCAGACGAAATTTATTTCAGAGCTGA	311	
Db	86	TTGAGGCCAAGGAAGAACCAACATATATGAGGTTGCCGATGAGATTTTTCGGAGCTGA	145	
Qy	312	AGTCCATGGCACATATTTGGTCAAGGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATG	371	
Db	146	AGTCCATTACGCAAGACGGTCTGGAGTTTGATGAGAAGAAATATTAGCGGAGGATATATG	205	
Qy	372	ATGCTTTCAAGGTTCTCATTTGCACTTTCGTTGTTATTGCAAAAGAAAGAAAGGACATCGGT	431	
Db	206	ATGCTTTTC-ATGTCCTCATTTGCAATTCGTTGTTATTGCAAAAGATAAAAGGATATAAAGT	264	
Qy	432	GGATGGGCGCTTTCCAAATTACAGATATCAAAAAATAAAGAGCTTTGAGGAAGTTTC-GTAAA	490	
Db	265	GGATGGGCGCTTACTAATTATAGATACGAAAGATACAGAAGTTTGGAGGAAGTTTCACAAA	324	
Qy	491	GAACCTGTCACAGATTAGGAACAAGAGGCACTCTCTCCAGGAATTCGAA-AAACAGTT	549	
Db	325	GAAATCATCACAGGATCCAGATTAAGAAAGAAAGCTTCTCCAGAAATTTGAATACGCAGAT	384	
Qy	550	TGATGATCTCCAAAACATCAAGTTTACG-TAACCAAACACTGGAAAGCTCAGCAGAGATG	608	
Db	385	TGATGACCTTCAGAAATTACATTAAGCAAAACACAGGCTAGTAGCGGCCAGGCAAAAGTG	444	
Qy	609	TTAATGGCATCC--GCTTTCCTATTCCGTATTGGTCAAGACA--TCTAGGAAGCAAGGGTG	664	
Db	445	TTAATGGGATTACATCTTACAGTTCTTAAATGATCAAGAAATTTCCGATATAAGCAAGGTTG	504	
Qy	665	GAAATTTGAGATTT	677	
Db	505	GAAATTCAAAATT	517	

RESULT 11  
BI802607

LOCUS	BI802607	548 bp	mRNA	linear	EST 01-NOV-2001
DEFINITION	H084D10 Endosperm library from <i>Oryza sativa</i> (10 days after anthesis); <i>Oryza sativa</i> cDNA clone H084D10, mRNA sequence.				
ACCESSION	BI802607				
VERSION	BI802607.1	GI:16575311			
KEYWORDS	EST.				
SOURCE	<i>Oryza sativa</i>				
ORGANISM	<i>Oryza sativa</i>				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; <i>Oryza</i> .				
REFERENCE	1 (bases 1 to 548)				

**AUTHORS** Dong, H.T., Li, D.B., Zhuang, X.F., Dai, C.G., Sun, L.X., Pei, Y.X.,  
 Wu, H.F., Jiang, Y.X., Yu, P.C., Gao, Q.K. and Lou, Y.C.

**TITLE** A Gene Expression Screen in *Oryza sativa*

**JOURNAL** Unpublished (2001)

**COMMENT** Contact: Haitao Dong, Debao Li  
 Bioinformatics and Gene Network Research Group  
 Zhejiang University  
 Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China  
 Tel: 0086-571-86892051  
 Fax: 0086-571-86961525  
 Email: webmaster@estarray.org, URL: http://www.estarray.org  
 Seq primer: M13 forward primer.

**FEATURES**  
 source  
 1..548  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4530"  
 /clone="H084D10"  
 /tissue\_type="Endosperm"  
 /dev\_stage="10 days after anthesis"  
 /clone\_lib="Endosperm library from *Oryza sativa* (10 days after anthesis)"  
 /note="Vector: pSport2"

**ORIGIN**  
 Query Match 20.0%; Score 218; DB 4; Length 548;  
 Best Local Similarity 71.1%; Pred. No. 5.6e-37;  
 Matches 391; Conservative 0; Mismatches 140; Indels 19; Gaps 7;

QY 554 GATCTCCAAAACATCAAGTTACGTACCAACCAACCTGAAAGCTCAGCAGAGAAATGTAAT 613  
 |||||  
 Db 1 GACCTTCAGAAATATACATATCCACACAGGCTAGTCAGAGGCCA-CAGAAAGTGTAAAT 59  
 |||||  
 QY 614 GGCAATCCGCTTCC-ATTCGTATGTGTCAGACATCTAGGAAAGAGGGTGAATTTGA 672  
 |||||  
 Db 60 GGCAATCCTCTCCGCTTCTATTGATCAAGACATCCGAAAGCAAGGGTGAATTTGA 119  
 |||||  
 QY 673 GATTCAGATGACTCCGAAGTTTCCCATTTCCAGTTTCAATGGTGCACCATTCACATTTGA 732  
 |||||  
 Db 120 GATTCGGAAGATTCGAAGTTTTCAGCGTTTCAGCTTCAACGGTGCACCATTCACCATGA 179  
 |||||  
 QY 733 TGATGATCTCTCAATCCTTGAGGGGTAAAGCGCTAACAGCATAGGAAGAGCTGCGCGCG 792  
 |||||  
 Db 180 TGATGATGATCAATCCTTGAAGCCATCAGGCGTAAACACAGGAAGAGCTGCGCTCTC 239  
 |||||  
 QY 793 CACCTTCTACT-AGAGCTCAAGAAATTTACAAATGAATTAAGTG-----TTAGAACTG 847  
 |||||  
 Db 240 CATTCACCTTTAAGAGGCCACAGAATGTGACAAACCATTTGAAGTGTCAAACTCAAACTG 299  
 |||||  
 QY 848 GCACAGCGGATCTTTTGACACAGCTATGTATAGCTATA-----TATCCTCATGAAA 900  
 |||||  
 Db 300 GCACACACAGTTCTTTTGACACAGTTATGTATAGCTATAGCGTACCTTTTCAGTATGAAA 359  
 |||||  
 QY 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGGCTTTGAGAAGA-TTTTAACTCGAAATTTT 959  
 |||||  
 Db 360 CTGCACTAGTTTATAGGACAGTCTCTCAGGCTTTTATAGCATATGACCTTGGAAATTT 419  
 |||||  
 QY 960 GTCTCCTTTTGTGCTTGTAGGCTTATAGGCTCTCAGATAGATGATTCATATATGTGCTG 1019  
 |||||  
 Db 420 GCCTCTTTTGTGCTTGTAGGCTGTA---GTCTGGAATAGTTTCTTGTGCTGTGCTG 475  
 |||||  
 QY 1020 CTATGAACATTTGATAGCAAAACAAAAACAAAAACAAAAACAAAAACAAAAACAAAA 1079  
 |||||  
 Db 476 TTATGTATAAATCAATTCGCGTTTCAATTCGAAGCTGTGTACAAGCTCCATTTGATGAAA 535  
 |||||  
 QY 1080 AAAAAAAAAA 1089  
 |||||  
 Db 536 AAAAAAAAAA 545  
 |||||

RESULT 12  
 CL909692/c 836 bp DNA linear GSS 14-SEP-2004  
 LOCUS

**DEFINITION** OA\_Aba0009114.f OA\_Aba *Oryza australiensis* genomic clone  
 OA\_Aba0009114 5', genomic survey sequence.

**ACCESSION** CL909692

**VERSION** CL909692.1 GI:52018571

**KEYWORDS** GSS.

**ORGANISM** *Oryza australiensis*  
*Oryza australiensis*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; *Oryza*.  
 1 (bases 1 to 836)  
 Kim, H., Yu, Y., Stum, D., Yost, D., Rao, K., Luo, M., Jetty, R.,  
 Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.  
 OMAP Project  
 Unpublished (2004)  
 Contact: Rod A. Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: http://genome.arizona.edu

**PCR Primers**  
 FORWARD: TAA TAC GAC TCA CTA TAG GG  
 BACKWARD: CAC TCA TTA GGC ACC CCA  
 Plate: 0009 row: 1 column: 14  
 Seq primer: TAA TAC GAC TCA CTA TAG GG  
 Class: BAC ends.

**Location/Qualifiers**  
 1..836  
 /organism="Oryza australiensis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4532"  
 /clone="OA\_Aba0009114"  
 /tissue\_type="young leaves"  
 /lab\_host="DH10B T1 phage resistant"  
 /clone\_lib="OA\_Aba"  
 /note="Vector: pGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

**ORIGIN**  
 Query Match 14.7%; Score 160.6; DB 9; Length 836;  
 Best Local Similarity 74.8%; Pred. No. 1.9e-24;  
 Matches 246; Conservative 0; Mismatches 69; Indels 14; Gaps 3;

QY 714 GTGCACCATTCACATTCATGATGATCTCTCAATCCTCAGGGGTAAAGCGTAACAGCA 773  
 |||||  
 Db 344 GTGCACCATTCACATTCATGATGATGATCTCTCAATCCTCAGGGGTAAAGCGTAACAGCA 285  
 |||||  
 QY 774 TAGGAAGAGCTGGCGCGCCACCTTCACT--AGAGACTCAAGAAATATTACAAATGAAT 831  
 |||||  
 Db 284 TAGGAAGAGCTGGCGCTCTCCATTCCTTTAAAGACACACAGAAATGTGACAAAACAGT 225  
 |||||  
 QY 832 AAAAGTGTAGAACTGGCAGCGGATTTCTTTTGACAGCTATGATAGCTAT----- 885  
 |||||  
 Db 224 TGAAGTGTCAAACTGGCAGTGGCAGTCTTTTGGACAGCTATGATAGCTATAGCGTA 165  
 |||||  
 QY 886 --ATATCCTCATGAAAACCTTGACCTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGATT 943  
 |||||  
 Db 164 CCATTCAGTAATGAAAACCTGACCTAGTTTATAGGACAGTATCTCAGGCTTTTAAACAAA 105  
 |||||  
 QY 944 TTAACCTGCAAAATTTGTCTCTTTTGTGCTAGCAGGTTATTAGTCTTCAGATAGATG 1003  
 |||||  
 Db 104 TATGACTCGGAATTTTGGCTCTTTTGTGCTTATAGCAGGTTGA-----GTCTGGATAGGTG 49  
 |||||  
 QY 1004 ATTATATATGCTGCTCTATGAAAACATT 1032  
 |||||  
 Db 48 TTCTTGCATGCTGCTGCTATGATAAACT 20  
 |||||

RESULT 13  
 CK278661 992 bp mRNA linear EST 03-AUG-2004  
 LOCUS  
 DEFINITION EST724739 potato abiotic stress cDNA library Solanum tuberosum cDNA

[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)  
<http://www.genoscope.cns.fr/cgi-bin/qqb/qqb?source=Arabidopsis>.

FEATURES		Location/Qualifiers	
source		1..1403	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/strain="Col-0"	
		/db_xref="taxon:3702"	
		/clone="GSLTSL672A12"	
		/tissue_type="Adult vegetative tissue"	
gene		/plasmid="pCMVSPORT_6"	
		1..1403	
		/gene="At5g03415"	
ORIGIN			
Query Match		13.8%;	Score 150.8; DB 3; Length 1403;
Best Local Similarity		55.3%;	Pred. No. 2.8e-22;
Matches	346; Conservative		0; Mismatches 262; Indels 18; Gaps 2;
QY	147	AGGGGCTGTGACCCCGGATAAGATAGGAAGAGGAGGCTGCGGACCGAGGATCA	206
Db	426	AAGAGCTTCTGTGTAAAGAAAGAGAGGGGACAGCGTGGCTGGTCCAGATAGA	485
QY	207	CCGGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAA	266
Db	486	CTGGAAGAGGACTACGTCATTTAGTATGAGAGTTTGTGAAAAGGTGGAAGCAAAGGAA	545
QY	267	GAACAATACATAGAGTTGCGAGCGAAATTTATTCAGAGCT-----GA	311
Db	546	GGACAACTTACAATGAGTTGCGAGCGAGCTTGTGCTGAATTTGCATCTTCCAATGACG	605
QY	312	AGTCCATGGCACATATTGGTCAAGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATG	371
Db	606	ATGGAACATCCCTGATCAGCAACATGATGATGAGAAACATAGACGAGAGTATATG	665
QY	372	ATGCTTTCAACGTTCTCATTCGACTTCTGTGTTATTCGAAAGAAAAAGGAGATACGOT	431
Db	666	ATGCTTTAAACGTTCTCATGCTATGATATAATACCAAGGATAAAAAAGAAATTCAT	725
QY	432	GGATGGCCCTTCAATACAGATATGAAAAATAAGAGCTTGGAGGATTCGTAAG	491
Db	726	GGAGAGGTCTTCTCGGACAAAGCTTAAGCGACATTAAGAATTAAGAAGCAACGACTCT	785
QY	492	AACCTGTCACAAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTG	551
Db	786	CATTAGGAACAGATTGAGAAGAAACTGCATATTTCCCAAGACTCGAAGACAATATG	845
QY	552	ATGATCTCCAAACATCAAGTTACGTAAACCAACATCTGGAAGACTCAGCAGAGAATGTTA	611
Db	846	TAGGCTTTCAGATCTGATACAGAGAAATGAGCACTTATATAGCTCAGGAAATGCTCCCA	905
QY	612	ATGCG---ATCCGCTTCCATTCGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	668
Db	906	GTGGCGGTGTGTCTTCTCTTTTATCTCTGTCAGACTCGTCTCAGCAACAGTAGAAG	965
QY	669	TTGAGATTTCAGATGACTCGAAAGTTTCCCAATTTCCAGTTTCAATGGTGCAACCATTCACAT	728
Db	966	TGAGATATCAGAGATATGACGCTGTGATTTGATTTCAACAGCACTCCATTTGAGC	1025
QY	729	TGATGATGATCTCTCAATCCTTGAAG	754
Db	1026	TCACAGCAGCAAAATTTTGTCTCTCAAG	1051
RESULT 15			
CNS09YMC		1418 bp	linear
LOCUS		Arabidopsis thaliana Full-length cDNA Complete sequence from clone	
DEFINITION		GSLTSL812A03 of Adult vegetative tissue of strain col-0 of	
		Arabidopsis thaliana (thale cress).	
ACCESSION		BX831357	
VERSION		BX831357.1	
KEYWORDS		HTC; GSLT cDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		AUTHORS	
1 (bases 1 to 1418)		Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.	
TITLE		Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1418)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)	
COMMENT		- Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_RF/Full length http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.	
FEATURES		Location/Qualifiers	
source		1..1418	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/strain="Col-0"	
		/db_xref="taxon:3702"	
		/clone="GSLTSL812A03"	
		/tissue_type="Adult vegetative tissue"	
		/plasmid="pCMVSPORT_6"	
		1..1418	
		/gene="At5g03415"	
ORIGIN			
Query Match		13.6%;	Score 148.2; DB 3; Length 1418;
Best Local Similarity		55.0%;	Pred. No. 1e-21;
Matches	345; Conservative		0; Mismatches 263; Indels 19; Gaps 2;
QY	147	AGGGGCTGTGACCCCGGATAAGATAGGAAGAGGAGGCTGCGGACCGAGGATCA	206
Db	410	AAGGAGCTTCTGTGTAAAGAAAGAGGAGGACAGCGTGGCTGGTCCAGATAAGA	469
QY	207	CCGGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAA	266
Db	470	CTGGAAGAGGACTACGTCATTTAGTATGAGAGTTTGTGAAAAGGTGGAAGCAAAGGAA	529
QY	267	GAACAATACATAGAGTTGCGAGCGAAATTTATTCAGAGCT-----GA	311
Db	530	GGACAACTTACAATGAGTTGCGAGCGAGCTTGTGCTGAATTTGCATCTTCCAATAACG	589
QY	312	AGTCCATGGCACATATTGGTCAAGGTTTGAAGAGAAATATTAGCGGAGAGTGTATG	371
Db	590	ATGGAACATCCCTGATCAGCAACATGATGATGAGAAAAACATAAGCAAGAGTATATG	649
QY	372	ATGCTTTCAACGTTCTCATTCGACTTCTGTGTTATTCGAAAGAAAAAGGAGATACGGT	431
Db	650	ATGCTTTAAACGTTCTCATGGCTATGATATATATATCCAAAGGATAAAAAAGAAATTCAT	709
QY	432	GGATGGCCCTTTCAAATTCAGATATGAAAAATAAGAGCTTGGAGGATTCGTAAG	491
Db	710	GGAGAGGTCTTCTCGGACAAAGCTTAAGCGACATTAAGAATTAAGAAGCAACGACTCT	769
QY	492	AACCTGTCACAAAGATTAGGAACAAGAGGCACTCTCTCCAGGAAATCGAAAAACAGTTTG	551

```

Db 770 CACTTAGGAACAGATTGAGAGAGAAACTGCATATATCCCAAGAACTGGAAGACAATATG 829
QY 552 ATGATCTCCAAAACATCAAGTTACGTAAACCAACACACTGGAAAGCTCAGCAGAGAAATGTTA 611
Db 830 TAGGCCTTCAGATCTGATACAGAGAAATGAGCACTTATATAGCTCAGGAAATGCTCCCA 889
QY 612 ATGGC----ATCGGCCTTCCATTCTGTTGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 667
Db 890 GTGGCGGTGTTGCTCTTCTCTTTTATCCTTTGTCAGACTCGTCTCAGCAACAGTAGAA 949
QY 668 ATTGAGATTTCAGATGACTCGAAGTTTGCCCATTTGAGTTCAATGGTGCACCAATTCACA 727
Db 950 GTGGAGATATCAGAGATATGCAGCTCGTGCATTTTGATTCAACAGCACTCCATTGGAG 1009
QY 728 TTGCATGATGATCTCTCAATCCTTGAG 754
Db 1010 CTCACGACGACAAATTTGTCCTCAAG 1036

```

Search completed: February 6, 2005, 15:46:10  
Job time : 4259 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:25:31 ; Search time 228 Seconds  
(without alignments)  
7815.374 Million cell updates/sec

Title: US-10-088-830-1  
Perfect score: 1089  
Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.8	9.1	2320	4	US-09-023-655-1456
2	97	8.9	1427	4	US-09-949-016-3349
3	86.2	7.9	1157	2	US-08-723-415B-5
4	86.2	7.9	1157	3	US-09-189-627A-5
5	86.2	7.9	1157	3	US-09-710-861-5
6	85	7.8	358	4	US-09-640-211A-566
7	83.2	7.6	1700	2	US-08-428-131-1
8	83.2	7.6	1700	3	US-08-602-846-1
9	83.2	7.6	1700	3	US-09-078-596-1
10	82.8	7.6	364	4	US-09-640-211A-465
11	81.2	7.5	1154	2	US-08-723-415B-3
12	81.2	7.5	1154	3	US-09-189-627A-3
13	81.2	7.5	1154	3	US-09-710-861-3
14	81.2	7.5	1202	3	US-08-723-415B-7
15	81.2	7.5	1202	3	US-09-189-627A-7
16	81.2	7.5	1202	3	US-09-710-861-7
17	81.2	7.5	1385	2	US-08-723-415B-1
18	81.2	7.5	1385	3	US-09-189-627A-1
19	81.2	7.5	1385	3	US-09-710-861-1
20	74.8	6.9	1440	4	US-09-023-655-1399
21	74.8	6.9	1672	4	US-09-949-016-2937
22	71.6	6.6	601	4	US-09-949-016-120497
23	71.6	6.6	601	4	US-09-949-016-120498
24	71.6	6.6	193169	4	US-09-949-016-15091
25	69.4	6.4	1461	3	US-08-722-126A-4
26	69.4	6.4	1461	5	PCT-US95-04258-4
27	69.4	6.4	7218	1	US-08-232-463-14

28	65.8	6.0	1008	4	US-09-780-641-1	Sequence 1, Appli
29	65.8	6.0	2202	3	US-09-465-558-59	Sequence 59, Appl
30	65.2	6.0	578	3	US-09-602-877A-95	Sequence 95, Appl
31	64.4	5.9	189	4	US-09-821-976-14761	Sequence 14761, A
32	63.8	5.9	1507	3	US-09-453-323-1	Sequence 1, Appli
33	63.8	5.9	1813	3	US-09-071-224-3	Sequence 3, Appli
34	63.6	5.8	1034	4	US-09-311-021-105	Sequence 105, App
35	63.6	5.8	1141	4	US-09-800-729-78	Sequence 78, Appl
36	63.2	5.8	1736	3	US-09-182-816-22	Sequence 22, Appl
37	63.2	5.8	1736	3	US-09-182-816-24	Sequence 24, Appl
38	63.2	5.8	1736	3	US-09-471-528-22	Sequence 22, Appl
39	63.2	5.8	1736	3	US-09-471-528-24	Sequence 24, Appl
40	63.2	5.8	1736	3	US-09-634-530-22	Sequence 22, Appl
41	63.2	5.8	1736	3	US-09-634-530-24	Sequence 24, Appl
42	62.8	5.8	1602	1	US-08-530-950-3	Sequence 3, Appli
43	62.8	5.8	1602	3	US-08-888-429A-3	Sequence 3, Appli
44	62.8	5.8	1602	3	US-09-149-879-3	Sequence 3, Appli
45	62.8	5.8	1602	4	US-09-057-009-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-023-655-1456  
; Sequence 1456, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; FILING DATE: HEREWITH  
; APPLICATION NUMBER: US/09/023,655  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1456:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2320 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: 9604478  
; US-09-023-655-1456

Query Match 9.1%; Score 98.8; DB 4; Length 2320;

Best Local Similarity 58.1%; Pred. No. 6.8e-15;  
Matches 200; Conservative 0; Mismatches 132; Indels 12; Gaps 1;  
QY 236 ATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGAGTTGACAGAGAA 295  
DB 369 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAGTCGCTGATGAG 428  
QY 296 ATTTATTCAGAGCTG-----AAGTCATGCGACATATTTGGTCAAGGGTTTGTAT 343  
DB 429 CTGGTGTGAGAGTTTCAACCAATTCAAATAACCAATTTGGCTGCTGATTCGCGAGGCTTAATGAT 488  
QY 344 GAGAAATATTTAGCGGAGAGTGTATGATGCTTTTCAAGCTTCTCATTCGACACTTCGTGTT 403  
DB 489 CAGAAACATTTAGCGGAGAGTTTATGATGCTTTAAATGTGCTATGCAATGAACATA 548  
QY 404 ATTGCAAAAGAAAAAGGAGATACCGTGTGATGGCTTTCAAAATACAGATATGAAAAA 463  
DB 549 ATTCAAAAGCAAAAAAGAAATCAAGTGTGATGGCTGCTACCAATTCGTCTCAGGAA 608  
QY 464 ATAAAGAGCTTCAGGAAGTTCTGAAGAACTGTCACAAAGATTAGGAACAGAGGCA 523  
DB 609 TGTCAAACTGTGAGATAGAGAGCGAGCGGATAGACGGATAAGCAGAAAGCGGGCC 668  
QY 524 CTCTCCAGGAATCGAAAAACAGTTTGTATGATCTCCAAAAACAT 567  
DB 669 CAGCTGCAAGAACTTCTCTACAGCAAACTGCTTTCAAAAACCT 712

RESULT 2  
US-09-949-016-3349  
; Sequence 3349, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3349  
; LENGTH: 1427  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3349

Query Match 8.9%; Score 97; DB 4; Length 1427;  
Best Local Similarity 57.8%; Pred. No. 1.6e-14;  
Matches 197; Conservative 0; Mismatches 135; Indels 9; Gaps 1;  
QY 236 ATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGAGTTGACAGAGAA 295  
DB 459 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAGTCGCTGATGAG 518  
QY 296 ATTTATTCAGAGCTG-----AAGTCATGCGACATATTTGGTCAAGGGTTTGTATGAG 346  
DB 519 CTGGTGTGAGAGTTTCAACCAATTCAAATAACCAATTTGGCTGCTGATTCGCGGCTTAATGATCAG 578  
QY 347 AAGAAATATTAGCGGAGAGTGTATGATGCTTTTCAAGCTTCTCATTCGACACTTCGTGTTATT 406  
DB 579 AAGAACATTAGCGGAGAGTTTATGATGCTTTAAATGTGCTAATGCAATGAACATAATT 638  
QY 407 GCAAAAGAAAAAGGAGATACCGTGTGATGGCTTTCAAAATACAGATATGAAAAATA 466  
DB 639 TCAAGGAAAAAAGAAATCAAGTGTGATTTGGCTGCTCTACCAATTCCTGCTCAGGAATGT 698

QY 467 AAGAAGCTTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTC 526  
DB 699 CAGATCTCGAGATAGAGAGCAGAGCGGATAGACGATTAAGCAGAGCGGCCAG 758  
QY 527 CTCGAGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACAT 567  
DB 759 CTGCAAGAACTTCTCTCTACAGCAAAATCGCTTTCAAAAAACCT 799  
RESULT 3  
US-08-723-415B-5  
; Sequence 5, Application US/08723415B  
; Patent No. 5859199  
; GENERAL INFORMATION:  
; APPLICANT: Lathangue, Nicholas B.  
; APPLICANT: delaluna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/723,415B  
; FILING DATE: 30-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9610195.1  
; FILING DATE: 15-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-220  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1157 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1110  
US-08-723-415B-5

Query Match 7.9%; Score 86.2; DB 2; Length 1157;  
Best Local Similarity 56.0%; Pred. No. 7.6e-12;  
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;  
QY 239 GTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGAGTTGACAGAGCAAAAT 298  
DB 184 GTGTGTGAGAAAGTTGAGCGGAAGGCAACAACCTTATCATAGAGTTAGTGATGAGCTG 243  
QY 299 TATTCAGAGCTGAAGTCCA-----TGGCACAATATTTGGTCAAGGGTTTGTATGAG 346  
DB 244 GTATCTGAGTTTACCAACTCAAAATAACCACTGATTCGCGAGGCTTATGATCAG 303  
QY 347 AAGAAATATTAGCGGAGAGTGTATGATGCTTTTCAAGCTTCTCATTCGACACTTCGTGTTATT 406  
DB 304 GAGAACATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGAACATAATT 363



QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTTCAAATTCAGATATGAAAAATA 466  
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423  
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTC 526  
Db 424 CAGAACTCGAAATCGAAGACGAGCGCGATAGACGGATAAAGCAGAGCGAGCCCGAG 483  
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567  
Db 484 CTACAGAACTTCCTCTCAGCAAAATGCTTTTAAAAAACCCT 524

RESULT 4  
US-09-189-627A-5  
; Sequence 5, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189,627A  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1110)  
US-09-189-627A-5

Query Match 7.9%; Score 86.2; DB 3; Length 1157;  
Best Local Similarity 56.0%; Pred. No. 7.6e-12;  
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;  
QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACAACTGAGGTTCGAGACGAAAT 298  
Db 184 GTGTGTGAGAAAGTTGAGCGGAAGGCAACACTTCATACATGAGGTGATGAGCTG 243  
QY 299 TATTGAGAGCTGAAGTCCA-----TGGCACATATTGGTCAAGGGTTTGATGAG 346  
Db 244 GTATCTGAGTTTACCAACTCAATAAACCATCTGGCAGCTGATTCGCGAGGCTTATGATCAG 303  
QY 347 AGAATATTAGCGGAGAGTGTATGATGCTTTCAAGCTTCATTGCACTTCGTGTTATT 406  
Db 304 GAGAACATTAGACGAGAGTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAAAT 363  
QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTTCAAATTCAGATATGAAAAATA 466  
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423  
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTC 526  
Db 424 CAGAACTCGAAATCGAAGACGAGCGCGATAGACGGATAAAGCAGAGCGAGCCCGAG 483  
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567  
Db 484 CTACAGAACTTCCTCTCAGCAAAATGCTTTTAAAAAACCCT 524

RESULT 5  
US-09-710-861-5  
; Sequence 5, Application US/09710861  
; Patent No. 6387649  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/710,861  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: US/09/189,627  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1110)  
US-09-710-861-5

; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/710,861  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: US/09/189,627  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1110)  
US-09-710-861-5  
Query Match 7.9%; Score 86.2; DB 3; Length 1157;  
Best Local Similarity 56.0%; Pred. No. 7.6e-12;  
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;  
QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACAACTGAGGTTCGAGACGAAAT 298  
Db 184 GTGTGTGAGAAAGTTGAGCGGAAGGCAACACTTCATACATGAGGTGATGAGCTG 243  
QY 299 TATTGAGAGCTGAAGTCCA-----TGGCACATATTGGTCAAGGGTTTGATGAG 346  
Db 244 GTATCTGAGTTTACCAACTCAATAAACCATCTGGCAGCTGATTCGCGAGGCTTATGATCAG 303  
QY 347 AGAATATTAGCGGAGAGTGTATGATGCTTTCAAGCTTCATTGCACTTCGTGTTATT 406  
Db 304 GAGAACATTAGACGAGAGTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAAAT 363  
QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTTCAAATTCAGATATGAAAAATA 466  
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423  
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTC 526  
Db 424 CAGAACTCGAAATCGAAGACGAGCGCGATAGACGGATAAAGCAGAGCGAGCCCGAG 483  
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567  
Db 484 CTACAGAACTTCCTCTCAGCAAAATGCTTTTAAAAAACCCT 524  
RESULT 6  
US-09-640-211A-566  
; Sequence 566, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021C1U  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 566  
; LENGTH: 358  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-640-211A-566



QY 202 GATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAA 261  
|||  
Db 384 GAAGAAATGGCAAGGGCTCGGCAATTCTCCATGAAGGTGTGTGAGAAAGTTGAGAGGAA 443  
QY 262 AGAAGAACACATACAAATGAGGTTGCGAGAGGAAATTTATTCAGAGCTGAAGTC----- 315  
|||  
Db 444 AGAACCACCTCTACAAATGAGGTGGTGCAGAGCTGGTGCAGAGTTTCAGCGCTGCCGA 503  
QY 316 -----CATGGCACATATTGGTCAAGGTTTGTATGAGAAATATTAGCGGAGAGTGTA 369  
|||  
Db 504 CAACCACTTCTACCAACGAATCAGCTTAATGACCAAGAAATCCGGCGGCTGTCTA 563  
QY 370 TGATGCTTCAAGCTTCTCATTCGCTTGTGTTTATTCGCAAAAGAAAAAGGAGATACG 429  
|||  
Db 564 CGATGCTTAAATGTCTAATGCCATGAACATCATCTCCAGGAGAGAGAGATCAA 623  
QY 430 GTGGATGGGCTTCAAAATACAGATATGAAAAAATAAGAAAGCTTGAGGAAGTTGCTAA 489  
|||  
Db 624 ATGGATCGGCTCCCAACCAACTCAGCTCAGGAGTGCAGAACTTAGAGGTGAGAGGCA 683  
QY 490 AGAATCGTCAACAGATTAGGACAGAGGAGGCACTCTCCAGGAATCGAAAAAGTT 549  
|||  
Db 684 GAGGAGGTGGAGAGGATCAACAGAGGAGTCTCAGCTCCAGGAGCTCATCTGCGAGCA 743  
QY 550 TGATGATCTCCAAACAT 567  
|||  
Db 744 AATTGCTTCAAGAACTT 761

RESULT 9  
US-09-078-596-1  
; Sequence 1, Application US/09078596  
; Patent No. 6150116  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,596  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1700 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS

LOCATION: 55..1284  
US-09-078-596-1  
Query Match 7.6%; Score 83.2; DB 3; Length 1700;  
Best Local Similarity 53.7%; Pred. No. 5.1e-11;  
Matches 203; Conservative 0; Mismatches 163; Indels 12; Gaps 1;  
QY 202 GATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAA 261  
|||  
Db 384 GAAGAAATGGCAAGGGCTCGGCAATTCTCCATGAAGGTGTGTGAGAAAGTTGAGAGGAA 443  
QY 262 AGAAGAACACATACAAATGAGGTTGCGAGAGGAAATTTATTCAGAGCTGAAGTC----- 315  
|||  
Db 444 AGAACCACCTCTACAAATGAGGTGGTGCAGAGCTGGTGCAGAGTTTCAGCGCTGCCGA 503  
QY 316 -----CATGGCACATATTGGTCAAGGTTTGTATGAGAAATATTAGCGGAGAGTGTA 369  
|||  
Db 504 CAACCACTTCTACCAACGAATCAGCTTAATGACCAAGAAATCCGGCGGCTGTCTA 563  
QY 370 TGATGCTTCAAGCTTCTCATTCGCTTGTGTTTATTCGCAAAAGAAAAAGGAGATACG 429  
|||  
Db 564 CGATGCTTAAATGTCTAATGCCATGAACATCATCTCCAGGAGAGAGAGATCAA 623  
QY 430 GTGGATGGGCTTCAAAATACAGATATGAAAAAATAAGAAAGCTTGAGGAAGTTGCTAA 489  
|||  
Db 624 ATGGATCGGCTCCCAACCAACTCAGCTCAGGAGTGCAGAACTTAGAGGTGAGAGGCA 683  
QY 490 AGAATCGTCAACAGATTAGGACAGAGGAGGCACTCTCCAGGAATCGAAAAAGTT 549  
|||  
Db 684 GAGGAGGTGGAGAGGATCAACAGAGGAGTCTCAGCTCCAGGAGCTCATCTGCGAGCA 743  
QY 550 TGATGATCTCCAAACAT 567  
|||  
Db 744 AATTGCTTCAAGAACTT 761

RESULT 10  
US-09-640-211A-465  
; Sequence 465, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; TITLE OF INVENTION: Modification of Gene Transcription  
; FILE REFERENCE: 11000.1021C1U  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 465  
; LENGTH: 364  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-640-211A-465  
Query Match 7.6%; Score 82.8; DB 4; Length 364;  
Best Local Similarity 57.9%; Pred. No. 3.4e-11;  
Matches 175; Conservative 0; Mismatches 112; Indels 15; Gaps 1;  
QY 209 GGTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAGA 268  
|||  
Db 57 GGGAGGGGACTTCGCCAGTTTCAGCATGAAAGTATGTCAAAGGTCGAGAGCAAGGGTCGG 116  
QY 269 ACAACATACAAATGAGGTTCAGACGAAATTTATTCAGAGCTGAA-----G 313  
|||  
Db 117 ACCAGTATATGAGGTTCAGATGAATTAGTTGAGAGATATGCAATCTTAACAGTCG 176  
QY 314 TCCATGGCACATATTGGTCAAGGTTTGTATGAGAAAGAAATATTAGCGGAGAGTTGATGAT 373  
|||  
Db 177 CTCAATTTCTCCTGATCAGCAACAATATGATGAGAAAAAATAGGCGGAGGGTGTATGAT 236

QY 374 GCTTTCAAGCTTCTCATTGCACTTCGTGTTATTTGCAAAAGAAAAAGAGATACGGTGG 433  
Db 237 GCATTGAATGTACTGATGGCAATGGGACATCATATCAAGGACAGGAAGGAAATTCAGTGG 296  
QY 434 ATGGGCTTTCAAAATACAGATATGAAAAATAAGAAAGCTTCAGGAAGTTCCGTAAGAA 493  
Db 297 AAGGGTTACCTAGCAGCAAGTCTATGACCTTGAAGACTTGAAGCAAGCCATGGGA 356  
QY 494 CT 495  
Db 357 TT 358

RESULT 11  
US-08-723-415B-3  
; Sequence 3, Application US/08723415B  
; Patent No. 5859199  
; GENERAL INFORMATION:  
; APPLICANT: Lathangue, Nicholas B.  
; APPLICANT: delaluna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/723,415B  
; FILING DATE: 30-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9610195.1  
; FILING DATE: 15-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-220  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1154 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1107  
US-08-723-415B-3

Query Match 7.5%; Score 81.2; DB 2; Length 1154;  
Best Local Similarity 55.0%; Pred. No. 1.4e-10;  
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;  
QY 239 GTTTGTGAGAAAGTTGAAGCAAGGAAGAACACATACAAATCAGGTTGCGAGCAAAATT 298  
Db 184 GTGTGTGAGAAAGTTGAGGAAAGGCAACACTTCATACAAATGAGGTGATGAGCTG 243  
QY 299 TATTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGTTTGTATGAGAAG 349  
Db 244 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGTGATTCCGGCTTATGATCAGGAG 303

QY 350 AATATTAGCGGAGAGTGTATGATGCTTTCAAGGTTCTCATTGCACCTTCGTGTTATTGCA 409  
Db 304 AACATTAGACGAAGAGTTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAATTTC 363  
QY 410 AAAGAAAAAAGGAGATACGGTGGATGGCCCTTTCAAAATTACAGATATGAAAAATAAAG 469  
Db 364 AAGAAAAAAGAAATCAAGTGGATTGGCTGCTTACCAATTTCTGCTCAGGATGCCAG 423  
QY 470 AAGCTTTGAGAAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTC 529  
Db 424 AACCTGAAATCGAGAAAGCAGAGCGGATAGAACGGATAAAGCAGAGCGAGCCAGCTA 483  
QY 530 CAGGAATCGAAAAACAGTTTGATGATCTCCAAAAACAT 567  
Db 484 CAAGAACTTCTCTCTTCAGCAAAATTGCTTTTAAAAAACCT 521

RESULT 12  
US-09-189-627A-3  
; Sequence 3, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189,627A  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1154  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1107)  
US-09-189-627A-3

Query Match 7.5%; Score 81.2; DB 3; Length 1154;  
Best Local Similarity 55.0%; Pred. No. 1.4e-10;  
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;  
QY 239 GTTTGTGAGAAAGTTGAAGCAAGGAAGAACACATACAAATCAGGTTGCGAGCAAAATT 298  
Db 184 GTGTGTGAGAAAGTTGAGGAAAGGCAACACTTCATACAAATGAGGTGATGAGCTG 243  
QY 299 TATTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGTTTGTATGAGAAG 349  
Db 244 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGTGATTCCGGCTTATGATCAGGAG 303  
QY 350 AATATTAGCGGAGAGTGTATGATGCTTTCAAGGTTCTCATTGCACCTTCGTGTTATTGCA 409  
Db 304 AACATTAGACGAAGAGTTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAATTTC 363  
QY 410 AAAGAAAAAAGGAGATACGGTGGATGGCCCTTTCAAAATTACAGATATGAAAAATAAAG 469  
Db 364 AAGAAAAAAGAAATCAAGTGGATTGGCTGCTTACCAATTTCTGCTCAGGATGCCAG 423  
QY 470 AAGCTTTGAGAAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTC 529  
Db 424 AACCTGAAATCGAGAAAGCAGAGCGGATAGAACGGATAAAGCAGAGCGAGCCAGCTA 483  
QY 530 CAGGAATCGAAAAACAGTTTGATGATCTCCAAAAACAT 567  
Db 484 CAAGAACTTCTCTCTTCAGCAAAATTGCTTTTAAAAAACCT 521

RESULT 13  
US-09-710-861-3  
; Sequence 3, Application US/09710861  
; Patent No. 6387649  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/710,861  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: US/09/189,627  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1154  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1107)  
US-09-710-861-3

Query Match 7.5%; Score 81.2; DB 3; Length 1154;  
Best Local Similarity 55.0%; Pred. No. 1.4e-10;  
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;  
  
QY 239 GTTTGTGAGAAAGTTTGAAGCCAAAGGAAGAAACAAACATACATGAGTTTCAGACGAAATT 298  
DB 184 GTGTGTGAGAAAGTTTGAAGCCAAAGGAAGAAACAACTTACATGAGTTAGTGTGAGCTG 243  
  
QY 299 TATTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGGTTTGTATGAGAAG 349  
DB 244 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 303  
  
QY 350 ATATTTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTCGACTTCGTGTATTGCA 409  
DB 304 AACATTAGACGAGAGTTTATGATGCTTTTAAATGTACTAATGGCGATGAACATAATTCA 363  
  
QY 410 AAAGAAAAAAGAGATACGCTGATGGGCTTTTCAAAATTTACAGATATGAATAAATAAG 469  
DB 364 AAGAAAAAAGAAATCAAGTGATGGCTTCCCTACCAATTCCTCAGGAATGCCAG 423  
  
QY 470 AAGCTTGAGAAAGTTCTGTAAGAACTCGTCAACAGATTAGGAACAAAGAGGCATCTCTC 529  
DB 424 AACCTGGAATCGAAGACGAGCGGGGATAGAACGGATAGAACGAGCGAGCCAGCTA 483  
  
QY 530 CAGGAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567  
DB 484 CAAGAACTTCTCTTCAGCAAAATTTGCTTTTAAAAACCT 521

RESULT 14  
US-08-723-415B-7  
; Sequence 7, Application US/08723415B  
; Patent No. 5859199  
; GENERAL INFORMATION:  
; APPLICANT: Lathangue, Nicholas B.  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA

ZIP: 22201-4741  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/723,415B  
; FILING DATE: 30-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9610195.1  
; FILING DATE: 15-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-220  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1202 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1155  
US-08-723-415B-7

Query Match 7.5%; Score 81.2; DB 2; Length 1202;  
Best Local Similarity 55.0%; Pred. No. 1.4e-10;  
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;  
  
QY 239 GTTTGTGAGAAAGTTTGAAGCCAAAGGAAGAAACAAACATACATGAGTTTCAGACGAAATT 298  
DB 232 GTGTGTGAGAAAGTTTGAAGCCAAAGGAAGAAACAACTTATACATGAGTTAGTGTGAGCTG 291  
  
QY 299 TATTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGGTTTGTATGAGAAG 349  
DB 292 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 351  
  
QY 350 ATATTTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTCGACTTCGTGTATTGCA 409  
DB 352 AACATTAGACGAGAGTTTATGATGCTTTTAAATGTACTAATGGCGATGAACATAAATTCA 411  
  
QY 410 AAAGAAAAAAGAGATACGCTGATGGGCTTTTCAAAATTTACAGATATGAATAAATAAG 469  
DB 412 AAGAAAAAAGAAATCAAGTGATGGCTTCCCTACCAATTCCTCAGGAATGCCAG 471  
  
QY 470 AAGCTTGAGAAAGTTCTGTAAGAACTCGTCAACAGATTAGGAACAAAGAGGCATCTCTC 529  
DB 472 AACCTGGAATCGAAGACGAGCGGGGATAGAACGGATAGAACGAGCGAGCCAGCTA 531  
  
QY 530 CAGGAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567  
DB 532 CAAGAACTTCTCTTCAGCAAAATTTGCTTTTAAAAACCT 569

RESULT 15  
US-09-189-627A-7  
; Sequence 7, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189,627A  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415

```
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
US-09-189-627A-7

Query Match          7.5%; Score 81.2; DB 3; Length 1202;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACATGAGGTTGCAGACGAAATT 298
Db |||||||
QY 232 GTGTGTGAGAAAGTTCAAGCGGAAAGGCACAACTTCATACATGAGGTAGCTGATGAGCTG 291
Db |||||||
QY 299 TATTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGGTTTGATGAGAAG 349
Db |||||||
QY 292 GTATCTGAGTTTACCACCTCAATATACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 351
Db |||||||
QY 350 AATATTAGCGGAGAGTGATGATGCTTTCAAGCTTCTCATTTGCCACTTCGTGTTATTGCA 409
Db |||||||
QY 352 AACATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGAACATAATTTC 411
Db |||||||
QY 410 AAAGAAAAGAGAGATACGGTGGATGGCGCTTTCAAATTACAGATATGAATAAATAAG 469
Db |||||||
QY 412 AAGGAAAAAAGAAATCAAGTGGATGGCGCTGCCCTACCAATTCGTCTAGGAATGCCAG 471
Db |||||||
QY 470 AAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACCTCCTC 529
Db |||||||
QY 472 AACCTGGAATCGAAGCAGAGCGCGGATAGACGGATTAAGCAGAGCGAGGCCAGCTA 531
Db |||||||
QY 530 CAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAAT 567
Db |||||||
QY 532 CAAGAACTTCTCCTTCAGCAAAATTGCTTTTAAAAACCT 569
Db |||||||
```

Search completed: February 6, 2005, 13:09:14  
Job time : 230 secs

854 AACGATGAAGTCTATACAAGCTCAGGAGAGCTTATAGTTGCCGATGAGATTTTTCGGAG 795

QY	308	CTGAAGTCCATGGGCACATATTGGTCAAGGGTTTGATGAGAAGAAATATTAGGCGGAGAGTG	367
DB	794	CTGAAGTCCATATTACGCGAAGCGGTCTGGAGTTTGATGAGAAGAAATATTAGGCGGAGGGA	735
QY	368	TATGATGCTTTCAAGGTTCTCATTTGCACCTTCGTGTTATTGCAAAAAGAAAAAGAGATA	427
DB	734	TATGATGCTTTCAATGTGCTCATTTGCAATTCGTGTTATTGCAAAAAGATAAATAAGAGATA	675
QY	428	CGGTGGATGGGCCCTTTCAAAATTA CAGATATGAAAAATAAAGAAAGCTTGAGAAAGTTGCT	487
DB	674	AAAGTGGATGGGCCCTTACTTAATTAGATACGAAAAAGATACGAAGTTGAGGAAAGTTCCAC	615
QY	488	AAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACCTCTCCAGGAAATCGAAAAACAG	547
DB	614	AAAGAACTCATCA CAGGATCAGANTAAAGAAGAAGCTTCTCCAGGAAATTTGAAAAGCAG	555
QY	548	TTTGATGATCTCCAAAAACATCAAGTTACGTTAA CCAAAACA CTGGAAGCTCAGCAGAGAT	607
DB	554	TTTGATGACCTTCAGATATTTACATTTACGCAACCAAGGCTAGTCAGAGGCCAGCAAAAGT	495
QY	608	GTTTAATGGCATCCGCTTTCCATTCGTATTTGTCAGACATCTAGGAAAGCAAGGCTGGAA	667
DB	494	GTTTAATGGCATCCCTTTCCGTTCTTATTTGATCAAGACATCCGAAAAAGCAAGGGTGGAA	435
QY	668	ATTGAGATTTCAGATGACTCGAAGTTTGCCCATTTTCGAGTTCAATTCGATGTCACATTCA	727
DB	434	ATTGAGATTTCGGAAGATTCTGAAAGTTTGCA CGGTTTCGACTTTCAACGGTGCACCATTCACC	375
QY	728	TTGCATGATGATCTCTCAATCCTTTGAGGGGTGAGGCGTTAACAGCATAGGAGAGAGCTGCG	787
DB	374	ATGCATGATGATGATCAATCCTTTGAAGCCATCAGGCGTTAACCAAAAGGAAGAGCTGCG	315
QY	788	CGCGCCACCCCTTCACT-AGAGACTCAAGAATATTACAAA-----TGAAATTAAGTGTAG	842
DB	314	CTCTCCATTCACCTTAAGAGGCA CAGAATGTGACAAACCATTTGAAGTGCAACTCA	255
QY	843	AACTGGCAGACCGGATCTTTTGGCAGACTATGATATAGCTATA-----TATCCTCAT	895
DB	254	AACTGGCACCAACCAAGTTTCTTTTGACAGTTATGATATAGCTATATAGCGTACCTTTTCAGTAT	195
QY	896	GAAAACTTGACTAGTTTATATAGGACAGTCTCTCAGGCTTGAGAAGA-TTTTAACTCGCAA	954
DB	194	GGAACCTCGACCTAGTTTATAGGACAGTCTCTCAGGCTTTTATGACAGATATGACCTTGGCA	135
QY	955	ATTTTGTCTCTTTTGTGCGCTAGCAGGTTATAGGTTCTCAGATAGATGATTCATATATG	1014
DB	134	ATTTTGTCTCTTTTGTGCGCTAGCAGGTGA-----GTCTGGAATAGTTTTTTTCTTTCGCGTG	79
QY	1015	TGCTGCTATGAAAAACATTTAGTACAAAAAATAAAAAAAAAAAAAAAAAAAAAA	1074
DB	78	TGCTGTTATGATATAAATCAATTGCGTTTCAATTCGAAGCTTGTGTACAGCTCCATTTGA	19
QY	1075	AAAAAAAAAAAAAAAA 1089	
DB	18	TGAAAAAAAAAAAA 4	

```

RESULT 2
US-10-425-115-4115
; Sequence 4115, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

```

```

: SEQ ID NO 4115
:
: LENGTH: 774
:
: TYPE: DNA
: ORGANISM: Zea mays
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: MRT4577_103754C.1
US_10-425-115-4115

```

Query Match	35.4%;	Score 385.4;	DB 18;	Length 774;
Best Local Similarity	74.0%;	Pred. No. 3.2e-82;		
Matches 547;	Conservative 0;	Mismatches 176;	Indels 16;	Gaps 4
QY	53	GCTACCGCCGACCTGGACCTTGACCGCGGTGCACATCTTCGAAGCTTCCAGTGTCCCGCCG	112	
Db	18	GCTGCGCGGAGCCGAGCTTACACAGCTTGGCATCAGGAGAGAGCTGGGATCCCTCG	77	
QY	113	CTTCCCGAA-----GGGGCGGTAAATGCGGTCCAAAGGAAGGGCGCTGTTGAACCG	163	
Db	78	CTTCCCGGACGGCGAAGCGTCGGCGGCAATGCAGCTGAGGGAGAAGAGAGGGTGATAAG	137	
QY	164	GATAAGATAGGAAGAGAGAGGCTCGCGCACCGAGGATCACGGTTCGGGGCTCGC	223	
Db	138	AACAAGAGCGAGGAAGGGTAAAGCTGGTACGCGAGCGGATCGCGGGTGGGGCTCGC	197	
QY	224	GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAAGCCAAAGGAAAGAAACAATACAATGAG	283	
Db	198	GAGTTCAGCAAGATAGTTTCTAAAAGGTTGAGACCAAGACGGAATCTTACAGTGA	257	
QY	284	GTTGCGAGCGAAATTTAATCAGAGCTGAAATGCCATATGGCACATATTTGTTCAAGGGTTGAT	343	
Db	258	GTTGCAGATGAAATTTGTGGAGAGCTGAAATG---TGACTCTTATTGGTCAAGAGTTTGA	314	
QY	344	GAGAGGAATATTAGGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTTGCATCTCGTCT	403	
Db	315	GAGAAGAAATTCAGGAGGAGAGTGTATGATGCTTTTAAATGTGCTAAATTCGACTACGTGT	374	
QY	404	ATTGCAAAAGAAAAGGAGATACGGTGGATGGCGCTTTCAAATTTACAGATATGAAAAA	463	
Db	375	ATCAAAAAGATGAAAAGAGATAAATGGATGGCTTGTCTAATTTCCAAATATGAAAAG	434	
QY	464	ATAAGAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGCA	523	
Db	435	AT---TCAGTTTGGAGGAACTCGCAAAAGAACTCATGTACAGGATTAAGAAACAAGAAAA	491	
QY	524	CTCTCCAGAAATCGAAA--AACAGTTTGTATGATCTCCAAACATCAAGTTTACGTAAACA	582	
Db	492	CTTCTCAGGAAATGAAAGACCAGCGTGAATGACCTCAGAAATCAAGTTTCCGCAACA	551	
QY	583	AACACTGGAAAGCTCAGCAGAGAATGTTAATGGCATCCGCTTCCATTCGTATTGGTCAA	642	
Db	552	GTTTCTACAGGCCACAGAGAGTGTGAATGGTATCTGCGCTTCCATTTCTTATTGGTCAA	611	
QY	643	GACATCTAGAAAGCAAGGTTGGAAATTTGAGATTTTCAGATGACCTCGAAGTTTGGCCATT	702	
Db	612	GGCATCAGAAAAAGCAAGGTTGGAGATTGAGATCTCAGAAACCTCAAAGTTTGTGGTTT	671	
QY	703	CGAGTTTCAATGGTGCACCATTCATTGTCATGATGATCTCTCAATCTTTCAGGGGGTAA	762	
Db	672	GGACTTCAACTGGTACCATTCACCTTGCAGCAGATGATCAATCTTTCAGGCCATCAG	731	
QY	763	GCCTAACAGCATAGGAAGA	781	
Db	732	GTGTAAATAGCGTGAATAA	750	

RESULT 3  
US-10-767-701-17642  
; Sequence 17642, Application US/10767701  
; Publication No. US200401726841  
; GENERAL INFORMATION: David K.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Chou, Yongwei





QY 332 -----CAAGGGTTTGATGAGAAATATATTAGCGGA 362  
Db 254 AGGCACCAGATCTCTGATAACCCCTAAACCGGCAACAATATGATGAGAAAATATACGACGA 313  
QY 363 GAGTGTATGATGCTTTCAACCGTTCTCATTTGACCTTCGTTGTTATTCGCAAAAAGAAAAGG 422  
Db 314 GAGTTTATGATGCTTTGAATGTTCTGATGGCTATGGACATTTATATCTAAAGATATAAAGG 373  
QY 423 AGATACGGTGGATGGCGCTTTCAAAATACAGATATGAAAAATAAAGAAAGCTTGGGAAG 482  
Db 374 AGATCCAGTGGAAAGGGCTTGCGCGTACTAGTATATAAGTGACATTTGAAGAAATGAAGCGG 433  
QY 483 TTCTGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAA 542  
Db 434 AGCTTGTGGGACTGAAAGGTAGGATTTGAAAGAAAAGTGTCTTACCTACAGAGCTACAG 493  
QY 543 AACAGTTTGTATGATCTCCAAAACATCAAGTTTACGTAACCAACACTGGAAGAGCTCAGAG 602  
Db 494 ACCAATATGATGCTGCAAAACCTGATTCACGAATATGAGCAATCATATGTTTTCAGGA 553  
QY 603 A---GAATGTTAAATGGCATCGCCCTTCATTCGTTATTTGCTCAAGACATCTAGGAAGCA 659  
Db 554 ACACGCTTCTGTGGAGTGGCTTTGCCATTTATCTAGTTTCAGACCCGACCTCATGCTA 613  
QY 660 GGGTGAATGAGATTTTCAGATGACTCGAAGTTTGGCCATTTTCGAGTTCAATGGTGCAC 719  
Db 614 CCGTGAAGTTGAGATATCAGAGATATGACAGTGTGTCATTTTTCATATAGCACCC 673  
QY 720 CATTCACATTCATGATGATCTCTCAATCTTTGAGGGGTAAAGGCTAACAGCATAGGA 779  
Db 674 CATTCGAGTGCACGACGACTCATACGCTCTAAAAGAAATGCGATTTCTGTGGAAGAGAAC 733  
QY 780 GAGCTGGCGCGC 792  
Db 734 AACATGACAGCAC 746

RESULT 7  
US-10-425-115-2033  
; Sequence 2033, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 2033  
; LENGTH: 1483  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_101857C.1  
US-10-425-115-2033

Query Match 12.2%; Score 133.4; DB 18; Length 1483;  
Best Local Similarity 51.4%; Pred. No. 1.4e-21;  
Matches 385; Conservative 0; Mismatches 331; Indels 33; Gaps 2;  
QY 24 CGCTCCCGCGGCGGAGCTGTCGCGCGCTACCGCGGCTACCGCGGCTGACCTGACCGGGTGC 83  
Db 111 CGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 170  
QY 84 ACATTTCTCGAAGCTTCCAGTGTCCCGCGTTCCTCCGAGCGGGGGGTATGCGGTCCAAA 143  
Db 171 GCGAGACACCGTCCGCGCGCGCGCTAAACGGAATTCACCCCGGCGGAGACGCTCCAT 230  
QY -144 GGAAGGGGGTGTGTGACCCCGGATTAAGATAGGAAGAGAGAGGCTGCGCGCACCGAGGA 203

QY 272 ACATCAATGAGTGTTCAGACGAAATTTATTCAGAGCTG-----AAGTCC 316  
Db 365 ACATCAATGAGTGTTCGAGTGTGCTGTAATTTCTGAACCAAGCAATAGTAA 424  
QY 317 ATGGCACATATTTGGTCAAGGGTTTGATGAGAAATATTAGCGGAGAGTGTATGATGCT 376  
Db 425 TTGCCCTCTGATCAGCAACAATATGATGAAAAACATCCGCGGAAGGCTTATGATGCT 484  
QY 377 TTCAAGTGTCTCATTCGACTTCGTTGTTATTCGAAAGAAAAGAGAGATACGGTGGATG 436  
Db 485 CTGAACGTTTCTCATGCGCAATGGAATATTTTCCAGGACAAAAGAAATTCATGGAGG 544  
QY 437 GGCCTTTCRAATTACAGATATGAAAAATAAGAACTTCGAGAAATTCGTTAAAGAACTC 496  
Db 545 GGCCTTCTCGTACTACTGATGATATTTGAGAACTAAAGACAGAGCGGCTTGGGCTC 604  
QY 497 GTCAACAAGATTAGGAACAGAGGCACTCTCCAGGAAATCGA-AAAACAGTTTGATGA 555  
Db 605 AGGAATAGAAATGAAAAGAAAACAGCTATCTGACGAGCTTGAGGGAGCAATTCATAGG 664  
QY 556 TCTCCAAACATCAAGTTAGTAAACCAACACTGGAAGCTCAGCAGAGAAATGTTAATGG 615  
Db 665 TCTTCAGAACCTTATTCAGCAACAGAGAGTATATAGCTCAAGAAATCCCTCCAGTGG 724  
QY 616 CA--TCGCGCTTCCATTCGTTATTTGGTCAAGACATCTAGGAACCAAGGGTGGAAATGAG 673  
Db 725 AAGTGATCTTACCTTTATTTGGTACAGACACGCGCTCATGCAACTGGGGGAATGGAA 784  
QY 674 ATTTCAGATGACTGAAAGTTGGCCATTTCCAGTTCAATGGTGCACCATTCATTTGCAT 733  
Db 785 ATATCAGAAGATATGACGCTGTGTTCAATTTGATTTCAATAGCACTCTCTTTTGAGCTGCAT 844  
QY 734 GATGATCTCTCAATCTCTGAGG 755  
Db 845 GACGACAAATATGTTCTCAAGG 866

RESULT 6  
US-10-425-115-87351  
; Sequence 87351, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 87351  
; LENGTH: 1585  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_179669C.1  
US-10-425-115-87351

Query Match 12.8%; Score 139.8; DB 18; Length 1585;  
Best Local Similarity 55.1%; Pred. No. 4.3e-23;  
Matches 338; Conservative 0; Mismatches 242; Indels 33; Gaps 2;  
QY 213 GGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAAACAA 272  
Db 134 GGGGACTCGGCAGTTTATGATGAAGTTTGTGAGAAAGTTGAAGTAAAGGGAGAACAA 193  
QY 273 CATACAATGAGTGTTCAGACGAAATTTATTCAGCTGAAGTCCATGCAATATGCT- 331  
Db 194 CATATAATGAGTGTGAGATGAACCTGTTGCTGAGTTTACAGACCCCAACAAATATATTG 253



```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3170-045-C12_FLI
US-10-425-114-20051

Query Match      11.4%; Score 124.6; DB 17; Length 1167;
Best Local Similarity 57.4%; Pred. No. 1.7e-19; Indels 3; Gaps 1;
Matches 245; Conservative 0; Mismatches 179;

QY 332 CAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 273 CAACAATATGATGAGAAAACATTCGTCGAAGGCTGTATGATGCTTTGAATGTTCTCATG 332
QY 392 GCACCTCGTGTTATGCAAAAGAAAAAGAGAGATACGGTGGATGGCCCTTCCAAATTAC 451
Db 333 GCAATGGATATTATTTCTAAGGATAAAAAGGAAATTCATGGAAGGCTCTCCCTGCTACT 392
QY 452 AGATATGAAAAATAAAGAGCTTTGAGGAGTTTCGTAAGAACTCGTCAACAGATTAGG 511
Db 393 AGTCTAAGCGATATTGAAGAGCTTAAAGTCGGAGCGCTTGGGCTAAGGAATAGAAITGAA 452
QY 512 AACAAAGAAAGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 453 AAGAAAGCAGCTTATTTGCAAGAGCTCGAGGAGCAATACATAGTCTTCAGAAACTTATA 512
QY 572 TTACGTAAACAAACACTGGAAGCTTGAAGAGCTTGAAGAGTTTCGTAAGAACTCGTCAACAGATTAGG 628
Db 513 CAACGAATAGCAACTATATAGCTCAGGAAATGCTCCCAATGGAGGTGTATCTTTGCC 572
QY 629 TTGCTATTGCTCAAGCATCTAGGAAGCAAGGTTGGAATTCAGATTTCAGATGACTCG 688
Db 573 TTCAATCTGTGAGACACACGGCCCATGCAACTGTTGAAGTGAATATCAGAAATATG 632
QY 689 AAGTTTGCCTATTCGAGTTCAATGTTGCGACCAATTCACATTCGATGATGATCTCTCAATC 748
Db 633 CAGCTTGTGATTTTGATTTCAATAGCACTCCCTTTGAGCTGATGATGATGATGATGTT 692
QY 749 CTTGAGG 755
Db 693 CTCGAGG 699

RESULT 10
US-10-424-599-43105
; Sequence 43105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43105
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_138923C.1
US-10-424-599-43105

Query Match      11.1%; Score 121.4; DB 17; Length 1751;
Best Local Similarity 56.9%; Pred. No. 1.2e-18; Indels 3; Gaps 1;
Matches 243; Conservative 0; Mismatches 181;

QY 332 CAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 715 CAACAATATGATGAGAAAAACATTCGTCGAAGGCTGTATGATGCTCTGAAATGTTCTCATG 774
```

```
QY 392 GCACCTTCGTGTTATTGCAAAAGAAAAAGAGATACGGTGGATGGGCTTTCAAATTAC 451
Db 775 GCAATGGATATTATTCTTAAGGCAAAAAGGAATTCATGGAAGGGTCTCCCTCGTACT 834
QY 452 AGATATGAAAAATAAAGAGCTTCGGAAGTTTCGTAAGAACTCGTCAACAGATTAGG 511
Db 835 AGTCTAAGCGATATTGAAGAGCTTAAAGTCAGAGCGCTCTTGGGCTCAGGAATAGAAATTGAA 894
QY 512 AACAAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 895 AAGAAAGCAGCCTATTTTGCAAGAGCTGGAGGAGCAATACATAGGCTCTTCAGAAACTTAT 954
QY 572 TTACGTAAACAAACACTGGAAGCTTCAGAGAGCTCAGCAGAGAAATGTTAAT---GGCATCCGCTTCCA 628
Db 955 CAACGAATAGCAACTATATAGCTCAGGAATGCTCCCAATGGAGGTGTATCTTTGCC 1014
QY 629 TTGCTATTGCTCAAGCATCTTAGGAAAGCAAGGTTGGAATTTGAGATTTCAGATGACTCG 688
Db 1015 TTATTCTGTCAGACACGCGCCCATGCAACTGTTGAAGTGAATAATCAGAAATATG 1074
QY 689 AAGTTTGCCTATTCGAGTTCAATGCTGACCACTTCACATTCGATGATGATCTCTCAATC 748
Db 1075 CAGCTTGTGCAATTTGATTTATTTTGGCTCCCTTTGAGATGCATGATGATAACTATGTT 1134
QY 749 CTTGAGG 755
Db 1135 CTCGAGG 1141

RESULT 11
US-10-425-114-36537
; Sequence 36537, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 36537
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE119B07_FLI
US-10-425-114-36537

Query Match      11.0%; Score 120; DB 17; Length 911;
Best Local Similarity 55.2%; Pred. No. 1.9e-18; Indels 3; Gaps 1;
Matches 256; Conservative 0; Mismatches 205;

QY 332 CAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 1 CAACAATATGATGAGAAAAATATACGACGAAGAGTTTATGATGCTTTGAATGTTCTGATG 60
QY 392 GCACCTTCGTGTTATTTCAAAAGAAAAAGAGATACGGTGGATGGGCTTTCAAATTAC 451
Db 61 GCTATGGACATTATATCTAAAGATAAAAAGAGATCCAGTGGAGGCTTCGCGGTACT 120
QY 452 AGATATGAAAAATAAAGAGCTTCGGAAGTTTCGTAAGAACTCGTCAACAGATTAGG 511
Db 121 AGTATAAGTGACATTCGAAGAAATGAAGACGAGCTTGTGGACTGAAGAGTAGGATTGAA 180
QY 512 AACAAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 181 AAGAAAGTGTCTTACCTCAGGAGCTTACAAAGACCAATATGTTAGTCTGCAAAAACCTGATT 240
```

QY 572 TTACGTAAACCAACACTGGAAGCTCAGCAGA---GAATGTTAATGTCATCCGCTTCCA 628  
Db 241 CAACGAAATGAGCAATCATATGTTTCAGGAACACGCTTCTGCTGAGTGGCTTTGCCA 300  
QY 629 TTGCTATTGTTCAAGACATCTAGGAAGCAAGGTGGAATTTGAGATTTGAGATGCTCG 688  
Db 301 TTTATCCTAGTTTCAGACCCGACCTCATGCTACTGTGGAAGTTGAAATATCAGAAGATATG 360  
QY 689 AAGTTTCCCATTTCCGAGTTCAATGTTGCACCTTACATTTGATGATGATCTCTCAATC 748  
Db 361 CAGCTGCTGATTTGACTTCAATAGCACCCTTCAGCTGCAATGAGTCTCATAGCTC 420  
QY 749 CTTGAGGGGTAAAGCGGTAAACAGCATAGGAGAGCTGCGCGCG 792  
Db 421 CTAAGAAGAAATGCGATTTCTGTGGAAGAGAACACATGACAGCAC 464

RESULT 12  
US-10-767-795-6106  
; Sequence 6106, Application US/10767795  
; Publication No. US2004018130A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10767,795  
; CURRENT FILING DATE: 2004-01-30  
; NUMBER OF SEQ ID NOS: 117596  
; SEQ ID NO 6106  
; LENGTH: 642  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C6102\_1  
US-10-767-795-6106

Query Match 9.4%; Score 102.8; DB 18; Length 642;  
Best Local Similarity 54.9%; Pred. No. 2.1e-14;  
Matches 225; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 332 CAAGGGTTTCATGAGAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTTCATT 391  
Db 85 CAACAGTTCAGAGAGAAACATCCGCCGAGGGTTTATGATGCTCTGATGTTCTAATG 144  
QY 392 GCACTTCGTGTTATTGCAAGAAAAAAGGAGATACGGTGGATGGCCCTTTCAAAATTAC 451  
Db 145 GCAATGGATATTATATCTAAGGATAAAAGAAATACAGGGAGGGACGGCCTCGTACT 204  
QY 452 AGATATGAAAAATAAGAGCTTGAGGAATGTCGTAAGAACTCGTCAACAAGATTAGG 511  
Db 205 GGCTGAGTGATATTGAAGATTAAAGACTGAACGGCTTGGGCTGAGAAATAGGATTGAT 284  
QY 512 AACAGAAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAACATCAAG 571  
Db 265 AAAAAAGCTGCCTATTGTGATGTAATTTGAGGAACAGTTTGTAGTCTTCAGAAATTTGAT 324  
QY 572 TTACGTAACCAACACTGGAAGCTCAGCAGAGAAT---GTTAATGGCATCCGCCCTTCCA 628  
Db 325 CAACGAAATGAGCAACTATACAGCTCAGGCAATGCTTCCGGTGGGGAGTGGCTTTACCT 384  
QY 629 TTGCTATTGTTCAAGACATCTAGGAAGCAAGGTGGAATTTGAGATTTGAGATGACTCG 688  
Db 385 TTTATCCTGTTGAGACACGCGCTCATGCAACTGTTGAGTAGAAATATCGGAAGATATG 444  
QY 689 AAGTTTGGCCATTTCCAGTTCAATGTTGAGGACCAATTCACATGTCATGATGA 738  
Db 445 CAGCTGCTGATTTTGACTTTAATAGCACTCCCTTTGAGCTCCACGATGA 494

RESULT 13  
US-10-198-846-10136  
; Sequence 10136, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10136  
; LENGTH: 2968  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1\_2  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-10136

Query Match 9.1%; Score 99.2; DB 14; Length 2968;  
Best Local Similarity 58.2%; Pred. No. 3.2e-13;  
Matches 199; Conservative 0; Mismatches 133; Indels 10; Gaps 1;

QY 236 ATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGGTTGAGAGCAA 295  
Db 403 AAAGTGTGTGAGAAAGTTTCAACGAAAGGTACACATCGTACAAATGAAGTGGTATGAG 462  
QY 296 ATTTATTTCAGAGCTGAAGT-----CCATGGCACATATTGGTCAAGGGTTTGATGA 345  
Db 463 CTGGTGTGAGAGTTTCAACATTAACATTTGGTGTCTGCTGATTCGGGCTTATGATCA 522  
QY 346 GAAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCATTCGCTGCTGTTAT 405  
Db 523 GAAGAAATATTAGCGGAGAGTGTATGATGCTTTAAATGTCTAATGCAATGAACATAAT 582  
QY 406 TGCAAAAGAAAAAAGGAGATACGGTGGATGGCCCTTTCAAATTCAGATATGAAAAAAT 465  
Db 583 TTCAAAGGAAAAAAGAAATCAAGTGGATTTGGCTTGCCTTACCCTTCTGCTCAGGAATG 642  
QY 466 AAAGAAAGCTTGGAGAGTTTCGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCAT 525  
Db 643 TCAGAAATCTGGAGATAGAAAGCAGAGCGGATAGAACGGATAGAACGATAGAACGAGCGGCCCA 702  
QY 526 CTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACAT 567  
Db 703 GCTGCAAGAACTTCTCTACAGCAAAATCGTTTCAAAAACCT 744

RESULT 14  
US-10-641-643-1456  
; Sequence 1456, Application US/10641643  
; Publication No. US20040077003A1  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
; TITLE OF INVENTION: GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO

STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1456:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2320 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: G604478  
SEQUENCE DESCRIPTION: SEQ ID NO: 1456 :  
US-10-641-643-1456

Query Match 9.1%; Score 98.8; DB 17; Length 2320;  
Best Local Similarity 58.1%; Pred. No. 3.6e-13;  
Matches 200; Conservative 0; Mismatches 132; Indels 12; Gaps 1;  
QY 236 ATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCGAGCGAA 295  
DB 369 AAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAAGTCGCTGATGAG 428  
QY 296 ATTATTTACAGAGCTG-----AAGTCCATGCCACATATTTGCTCAAGGGTTTGAT 343  
DB 429 CTGGTGTACAGAGTTTCAACATTAACATTTGGCTGCTGATTCGAGGCTTAATGAT 488  
QY 344 GAGAAAGAAATTTAGCGGAGAGTGTATGATGCTTTCAACGCTTCTCAATTCACATTCGCTGTT 403  
DB 489 CAGAAAGCAATTAGCGGAGAGTTTATGATGCTTTAAATGTGCTAATGGCAATGAACATA 548  
QY 404 ATTGCAAAAGAAAAAGGAGATACGTTGGATGGCCCTTTCAAAATTACAGATATGAAAAA 463  
DB 549 ATTTCAAAGGAAAAAAGAAATCAAGTGGATTGGCCCTGCTACCAATTCGCTCAGGAA 608  
QY 464 ATAAAGAGCTTGAGGAAGTTTCGTAAGAACTCGTCAACAGATTTAGGAACAAGAGGCA 523  
DB 609 TGTCAAGATCTGGAGATAGAGAACAGAGCGGAGATAGCGATTAAGCAGAGCGGCCC 668  
QY 524 CTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567  
DB 669 CAGCTGCAAGAACTTCTCCTACAGCAAAATCGCTTTCAAAAACCT 712

RESULT 15  
US-09-954-456-1176  
; Sequence 1176, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456

CURRENT FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1176  
LENGTH: 1266  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-456-1176

Query Match 8.9%; Score 97; DB 9; Length 1266;  
Best Local Similarity 57.8%; Pred. No. 7.3e-13;  
Matches 197; Conservative 0; Mismatches 135; Indels 9; Gaps 1;  
QY 236 ATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCGAGCGAA 295  
DB 261 AAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAAGTCGCTGATGAG 320  
QY 296 ATTATTTACAGAGCTG-----AAGTCCATGCCACATATTTGCTCAAGGGTTTGATGAG 346  
DB 321 CTGGTGTACAGAGTTTCAACATTAACATTTGGCTGCTGATTCGCTTATGATCAG 380  
QY 347 AAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAATGCACTTCGTTGTTATT 406  
DB 381 AAGAACATTTAGCGGAGAGTTTATGATGCTTTAAATGTCTAATGGCAATGAACATAATT 440  
QY 407 GCAAAAGAAAAAGGAGATACGTTGGATGGCCCTTTCAAAATTACAGATATGAAAAATA 466  
DB 441 TCAAGAGAAAAAAGAAATCAAGTGGATTGGCTGCTACCAATTTCTGCTCAGGAATGT 500  
QY 467 AAGAGCTTTGAGGAAGTTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTC 526  
DB 501 CAGATCTCGAGATAGAGAGCAGAGCGCGGATAGAACGATTAAGCAGAGCGGCGCCAG 560  
QY 527 CTCACGGAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567  
DB 561 CTGCAAGAACTTCTCCTACAGCAAAATCGCTTTCAAAAACCT 601

Search completed: February 6, 2005, 17:24:00  
Job time : 693 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 20:05:19 ; Search time 165 Seconds  
(without alignments)  
611.785 Million cell updates/sec

Title: US-10-088-830-2  
Perfect score: 1319  
Sequence: 1 MAPPRGGAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	99.6	261	4	AAB67762
2	890.5	67.5	294	8	ADQ36843
3	492	37.3	344	3	AAG22851
4	492	37.3	385	3	AAG22850
5	492	37.3	385	5	AU72542
6	492	37.3	385	8	ADO62143
7	492	37.3	385	8	ADO63083
8	471.5	35.7	346	5	AU72514
9	466.5	35.4	376	5	AU72514
10	466.5	35.4	379	8	ADQ36855
11	461	35.0	299	3	AAB07975
12	437	33.1	277	3	AAG22852
13	428.5	32.5	292	5	AU72529
14	428.5	32.5	292	6	ABP96848
15	428.5	32.5	292	7	ADH59560
16	428.5	32.5	292	8	ADO63085
17	428.5	32.5	292	8	ADO62145
18	428.5	32.5	296	5	AU72497
19	416.5	31.6	194	2	AAV32164
20	412	31.2	262	5	AAU72572
21	406.5	30.8	251	5	AU72559
22	399.5	30.3	431	8	ADN05730
23	396.5	30.1	424	4	ABG28057
24	389	29.5	446	2	AAV18026
25	378.5	28.7	410	2	AAR89212

## ALIGNMENTS

### RESULT 1

AAB67762  
ID AAB67762 standard; protein; 261 AA.

XX AC AAB67762;

DT XX (first entry)

DE 11-JUN-2001 (first entry)

XX Amino acid sequence of a wheat E2F-dimerisation partner (DP) protein.

KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;

XX S phase; cell cycle; retinoblastoma protein; alter cell proliferation.

OS Triticum monococcum.

XX WO200121644-A2.

XX 29-MAR-2001.

PF 25-SEP-2000; 2000WO-EP009325.

XX 24-SEP-1999; 99ES-00002127.

PR 11-NOV-1999; 99ES-00002474.

XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

PI Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/26.

DR N-PSDB; AAF80144.

XX New isolated, enriched, cell free and/or recombinant nucleic acid useful for e.g. altering cell proliferation characteristic such as to alter plant cell, organ or tissue size.

PT Claim 7; Fig 1; 77pp; English.

PS The present sequence represents a E2F-dimerisation partner (DP) protein.

XX The protein acts as a plant E2F transcription factor. E2F and DP are two proteins that hetero-dimerise to form an active transcription factor that regulates G1 to S phase of the cell cycle, and later, the expression of genes required for S-phase progression. E2F and retinoblastoma protein also interact as a hetero-dimer in cells to suppress certain genes. This repression involves binding of the retinoblastoma protein to the E2F-DP dimer that is in turn bound to sites on DNA through the E2F DNA binding domain. DP proteins can be modulated to alter plant cell, organ or tissue shape, and particularly to alter cell proliferation characteristic such as to alter plant cell, organ or tissue size

Abb60852 Drosophil  
Aag74072 Human col  
Adm78625 Human tra  
Ado19869 Human PRO  
Aar54663 Transcrip  
Aaw83389 Caenorhab  
Adp12553 Protein e  
Abr82450 Human ARP  
Adq74871 Human and  
Aab67766 Fragment  
Aay32161 Impatiens  
Aau72563 Arabidops  
Aab67767 Fragment  
Aab67765 Fragment  
Aab33108 Pinus rad  
Aau72569 Arabidops  
Aau72557 Arabidops  
Aab67768 Fragment  
Aab33007 Pinus rad  
Abg05510 Novel hum

26 375.5 28.5 445 4 AAB60852  
27 374.5 28.4 355 4 AAG74072  
28 374.5 28.4 410 8 ADM78625  
29 374.5 28.4 410 8 ADO19869  
30 372.5 28.2 410 2 AAR54663  
31 354 26.8 575 2 AAW83389  
32 302 22.9 345 8 ADP12553  
33 299.5 22.7 405 6 ABR82450  
34 299.5 22.7 405 8 ADQ74871  
35 286 21.7 56 4 AAB67766  
36 284.5 21.6 138 2 AAY32161  
37 284.5 21.6 193 5 AAU72563  
38 272 20.6 56 4 AAB67767  
39 264 20.0 51 4 AAB67765  
40 261.5 19.8 119 3 AAB33108  
41 261 19.8 142 5 AAU72569  
42 261 19.8 142 5 AAU72557  
43 256 19.4 50 4 AAB67768  
44 254.5 19.3 120 3 AAB33007  
45 245 18.6 149 4 ABG05510



XX SQ Sequence 261 AA;  
Query Match 99.6%; Score 1314; DB 4; Length 261;  
Best Local Similarity 99.6%; Pred. No. 5.6e-115;  
Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPEAGNAVORKGAVDPDKDRKKEKAAAP 60  
DB 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPERGNAVORKGAVDPDKDRKKEKAAAP 60  
QY 61 RITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKMAHIGQGFDEKNTIRRRVYDAF 120  
DB 61 RITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKMAHIGQGFDEKNTIRRRVYDAF 120  
QY 121 NVLIARLVIAREKKEIRWMLGSLNRYEKKLEVRKELVKNIRKALLQETEKQFDDL 180  
DB 121 NVLIARLVIAREKKEIRWMLGSLNRYEKKLEVRKELVKNIRKALLQETEKQFDDL 180  
QY 181 ONIKLRNQTLESSAENVNGIRLPVLVKTSRKARVEIEISDDSKFAHFEFNGAPFTLHDD 240  
DB 181 ONIKLRNQTLESSAENVNGIRLPVLVKTSRKARVEIEISDDSKFAHFEFNGAPFTLHDD 240  
QY 241 LSILEGVRNRSIGRAGRATLH 261  
DB 241 LSILEGVRNRSIGRAGRATLH 261  
RESULT 2  
ADQ36843  
ID ADQ36843 standard; protein; 294 AA.  
AC ADQ36843;  
XX  
DT 07-OCT-2004 (first entry)  
DE  
DE Os018989-4003 protein.  
XX cell proliferation related polypeptide; cell proliferation; senescence;  
KW differentiation; stress response.  
XX  
OS Oryza sativa.  
XX  
XX W02004061122-A2.  
XX  
PD 22-JUL-2004.  
XX  
XX 23-DEC-2003; 2003WO-US041200.  
XX  
XX 26-DEC-2002; 2002US-0436565P.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Cooper B;  
XX  
XX WPI; 2004-534388/51.  
XX  
PT New nucleic acid molecule encoding a cell proliferation-related  
PT polypeptide, useful for modulating cell proliferation, senescence,  
PT differentiation, development, and stress response in plants, and for  
PT producing enhanced food crops.  
XX  
PS Claim 28; SEQ ID NO 2; 409pp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a cell proliferation-related polypeptide. The nucleic acid  
CC molecule and the encoded polypeptide, and methods are useful for  
CC modulating cell proliferation, senescence, differentiation, development,  
CC and stress response in plants, and for producing enhanced food crops. The  
CC present sequence represents a cell proliferation-related polypeptide. The  
CC present sequence is published separately from the main body of the  
CC specification as EPO data.  
XX

SQ Sequence 294 AA;  
Query Match 67.5%; Score 890.5; DB 8; Length 294;  
Best Local Similarity 63.1%; Pred. No. 4.7e-75;  
Matches 185; Conservative 31; Mismatches 36; Indels 41; Gaps 4;  
QY 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPE----- 34  
DB 1 MAPPCDDAAAAASAAPGLANLIRGAGLPSEPERYPPRPCTSDSFAPISREGDDIPPQ 60  
QY 35 -----AGNAVORK--GAVDPDKDRKKEKAAAPRITGWLREYSKIVCEKVEAKGR 83  
DB 61 KKSVSILRSGGGGNAAREBEGA---NRNGKKEKTAQRIITGWLREFSKIVSKVEAKGR 117  
QY 84 TTYNEVADEIYSELKMAHIGQGFDEKNTIRRRVYDAFNVLIARLVIAREKKEIRWMLGSLN 143  
DB 118 TTYNEVADEIYSELKMAHIGQGFDEKNTIRRRVYDAFNVLIARLVIAREKKEIRWMLGSLN 177  
QY 144 YYEYKIKLEEVKELVKNIRKALLQETEKQFDDLQNIKLNRNQTLESSAENVNGIRLP 203  
DB 178 YYEYKIKLEEVKELVKNIRKALLQETEKQFDDLQNIKLNRNQTLESSAENVNGIRLP 237  
QY 204 FVLVKTSRKARVEIEISDDSKFAHFEFNGAPFTLHDDLSILEGVRNRSIGRAG 256  
DB 238 FLIIKTSRKARVEIEISDDSKFAHFEFNGAPFTLHDDLSILEGVRNRSIGRAG 289  
RESULT 3  
AAG22851  
ID AAG22851 standard; protein; 344 AA.  
AC AAG22851;  
XX  
DT 17-OCT-2000 (first entry)  
DE  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25936.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 29-MAR-1999; 99US-0126264P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 04-MAY-1999; 99US-0132407P.  
XX 05-MAY-1999; 99US-0132484P.  
XX 06-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 07-MAY-1999; 99US-0132487P.  
XX 11-MAY-1999; 99US-0132863P.  
XX 14-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.



PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 24-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 16-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.

```
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161202P.
PR 28-OCT-1999; 99US-0161192P.
PR 28-OCT-1999; 99US-0161193P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.3%; Score 492; DB 3; Length 344;
Best Local Similarity 47.7%; Pred. No. 1.8e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVORRGAVDPDKRKKAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADRIYSE 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 33 GDDAGSGAGGVKKKKRGQAAGPKTGRLGRQFSMKVCEKVESKGRTTYNEVADELVAE 92
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 97 LKSMAHIG-----QGFEKIRRRVYDAFNVLIARLVIKKEKKEIRWGLSNRYREKIKK 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 FALPNNDGTSPDQOQYDEKIRRRVYDALNVLMAMDIISKDKKEIQWRGLPRTSLSDIEE 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 152 LEEVRKELVNKIRKKALLQIEKQFDDLNKIKRLNOTLESSAENVN-GIRLPFVLVKTS 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 LKNERLSLRNRIKKKTAYSOLEEQVGLQNLQORNEHLYSSGNAPSGGVALPFILVQTR 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 211 RKARVEISDDSKFAHFENGAPFTLHDDLSTILEGVR 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 PHATVEVEISEDQLVHFDFNSTPFELHDDNDFVLKTMK 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
AAG22850
ID AAG22850 standard; protein; 385 AA.
XX
AC AAG22850;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25935.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 18-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
```

```
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145291P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.3%; Score 492; DB 3; Length 385;
Best Local Similarity 47.7%; Pred. No. 2.2e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQKGAVDPKDKKKKAAAPRITGWLREYSKIVCEKAEKGRITTYNEVADEITYSE 96
DB 74 GDDAGSQGASGVKKKRGQRAAGPKTGRGLKQFSMKVCEKESKGRITTYNEVADELVAE 133
QY 97 LKSMHIG----QGFEKNIRRRVYDAFNVLIALRVIAKEKKEIRWMGLSNRYRKIKK 151
DB 134 FALPNNDTGSPDQQQYDEKNIRRVYDALNVLMDIISKDKKEIQWRGLPRTSLSDIEE 193
QY 152 LEEVRKELVKNIRKKALLOEIEKQFDDDLQNLKRNQTLLESSAENVN-GIRLPFVLVKT 210
DB 194 LKNERLSLRNRIEKTAYSQELEEQVGLQNLQORNEHLYSSGNAPSGGVLPFILVQTR 253
QY 211 RKARVEIEISDDSKFAHFEFNGAPFTLHDDLSTLEGVR 248
DB 254 PHATVEVEISEDMQLVHFDFNSTPFELHDDNFVLTKMK 291

RESULT 5
AAU72542
ID AAU72542 standard; protein; 385 AA.
XX
AC AAU72542;
XX
DT 26-FEB-2002 (first entry)
XX
DE Arabidopsis cell cycle protein CCP33.
XX
KW Cell cycle protein; CCP; cell cycle regulation; herbicide;
KW plant growth regulator; plant development; abiotic stress; biotic stress;
KW nutrient deprivation; pathogen attack; crop yield; immunogen.
XX
OS Arabidopsis thaliana.
XX
PN WO200185946-A2.
XX
PD 15-NOV-2001.
XX
PF 14-MAY-2001; 2001WO-IB001307.
XX
PR 12-MAY-2000; 2000US-0204045P.
XX
PA (CROP-) CROPDESIGN NV.
PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
XX
DR WPI; 2002-062249/08.
DR N-PSDB; AAS96332.
XX
PT New cell cycle protein and nucleic acid molecule encoding it useful for
PT regulating cell cycle progression in plants and for identifying
PT modulators which are useful as herbicides or plant growth regulators.
XX
PS Claim 34; Fig 39; 316pp; English.
```



QY 97 LKSMHIG-----QGFDEKNIRRVYDAFNVLIALRVIAKEKEIRWMGLSNRYREKIKK 151  
 Db 134 FALPNDGTSPDQQYDEKNIRRVYDALNVLNMDIISKDKKEIQWRGLPRTSLSDIEE 193  
 QY 152 LEEVRKELVNKIRKALLQIEIKQFDDIQLNIKRLNQTLESSAENVN-GIRLPFVLVKT 210  
 Db 194 LKNERLSLRNRIEKTAYSQELBEQVGLQNLQIRNEHLYSSGNAPSGGVALPFILVQTR 253  
 QY 211 RKARVEIETSDSKFAHFEFNGAPFTLHDDLSTLEGVR 248  
 Db 254 PHATVEISEDMQLVHDFNSTPFELHDDNFVLKTMK 291

RESULT 7  
 ID ADO63083 standard; protein; 385 AA.  
 AC ADO63083;  
 DT 15-JUL-2004 (first entry)  
 DE Transcription factor G2981, SEQ ID 1550.  
 KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
 KW osmotic stress tolerance; cold tolerance; heat tolerance;  
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;  
 KW glyphosate resistance; flowering; fertility; seed development.  
 OS Arabidopsis thaliana.  
 XX WO2004031349-A2.  
 XX 15-APR-2004.  
 XX 18-SEP-2003; 2003WO-US030292.  
 XX 18-SEP-2002; 2002US-0411837P.  
 PR 17-DEC-2002; 2002US-0434166P.  
 PR 24-APR-2003; 2003US-0465809P.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL,  
 PI Riechmann JB, Haake V, Dubell AN, Keddle JS, Sherman BK;  
 XX WPI; 2004-330163/30.  
 DR N-PSDB; ADO63082.  
 XX New recombinant polynucleotide encoding transcription factor  
 PT polypeptides, useful for producing transgenic plants with advantageous  
 PT properties compared to a reference plant.  
 XX Claim 1; SEQ ID NO 1550; 510pp; English.

The present invention relates to novel plant transcription factor  
 CC proteins (I) and nucleotide sequences (II) (ADO61534-A063778). The  
 CC sequences can be used to produce transgenic plants, which overexpress  
 CC (II), where the transgenic plant has an altered trait as compared to a  
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises  
 CC an altered trait selected from increased tolerance to abiotic stress,  
 CC increased tolerance to osmotic stress, increased tolerance to cold,  
 CC increased germination in cold, increased tolerance to heat, increased  
 CC germination in heat, increased tolerance to freezing conditions,  
 CC increased tolerance to low nitrogen conditions, increased tolerance to  
 CC low phosphate conditions, increased tolerance to disease, including  
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
 CC increased tolerance to multiple fungal pathogens, increased resistance to  
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance  
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late  
 CC flowering, altered flower structure, loss of flower determinacy, reduced  
 CC fertility, altered shoot meristem development, altered branching pattern,

CC altered stem morphology, altered vascular tissue structure, reduced  
 CC apical dominance, altered trichome density, altered trichome development,  
 CC altered trichome structure, altered root development, altered shade  
 CC avoidance, altered seed development, altered seed ripening, altered seed  
 CC germination, slow growth, fast growth, altered cell differentiation,  
 CC altered cell proliferation, altered cell expansion, altered phase change,  
 CC altered senescence, abnormal embryo development, altered programmed cell  
 CC death, lethality when overexpressed, altered necrosis patterns, increased  
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark  
 CC green leaves, change in leaf shape, increased leaf size and mass, light  
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,  
 CC altered seed coloration, altered seed size, altered seed shape, large  
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil  
 CC content, altered seed protein content, altered seedprenyl content,  
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and  
 CC decreased anthocyanin levels. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 385 AA;

Query Match 37.3%; Score 492; DB 8; Length 385;  
 Best Local Similarity 47.7%; Pred. No. 2.2e-37;  
 Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQRKGAVDPDKRKKEKAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADIYSE 96  
 Db 74 GDDAGSQGASGVKKRGORAAQPKTGRGLQFQSMKVCESKGRRTTYNEVADSLVAE 133  
 QY 97 LKSMHIG-----QGFDEKNIRRVYDAFNVLIALRVIAKEKEIRWMGLSNRYREKIKK 151  
 Db 134 FALPNDGTSPDQQYDEKNIRRVYDALNVLNMDIISKDKKEIQWRGLPRTSLSDIEE 193  
 QY 152 LEEVRKELVNKIRKALLQIEIKQFDDIQLNIKRLNQTLESSAENVN-GIRLPFVLVKT 210  
 Db 194 LKNERLSLRNRIEKTAYSQELBEQVGLQNLQIRNEHLYSSGNAPSGGVALPFILVQTR 253  
 QY 211 RKARVEIETSDSKFAHFEFNGAPFTLHDDLSTLEGVR 248  
 Db 254 PHATVEISEDMQLVHDFNSTPFELHDDNFVLKTMK 291

RESULT 8  
 ID ADO36845 standard; protein; 346 AA.  
 XX ADQ36845;  
 AC ADQ36845;  
 DT 07-OCT-2004 (first entry)  
 XX Cell proliferation-related polypeptide #1.  
 DE Cell proliferation related polypeptide; cell proliferation; senescence;  
 XX differentiation; stress response.  
 KW Oryza sativa.  
 OS WO2004061122-A2.  
 XX 22-JUL-2004.  
 XX 23-DEC-2003; 2003WO-US041200.  
 PR 26-DEC-2002; 2002US-0436565P.  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Cooper B;  
 XX WPI; 2004-534388/51.  
 XX New nucleic acid molecule encoding a cell proliferation-related

PT polypeptide, useful for modulating cell proliferation, senescence,  
PT differentiation, development, and stress response in plants, and for  
XX producing enhanced food crops.  
XX  
XX Claim 28; SEQ ID NO 4; 408pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a cell proliferation-related polypeptide. The nucleic acid  
XX molecule and the encoded polypeptide, and methods are useful for  
XX modulating cell proliferation, senescence, differentiation, development,  
XX and stress response in plants, and for producing enhanced food crops. The  
XX present sequence represents a cell proliferation-related polypeptide. The  
XX present sequence is published separately from the main body of the  
XX specification as EPO data.  
XX  
XX Sequence 346 AA;  
  
Query Match 35.7%; Score 471.5; DB 8; Length 346;  
Best Local Similarity 39.6%; Pred. No. 1.6e-35;  
Matches 113; Conservative 44; Mismatches 81; Indels 47; Gaps 8;  
  
QY 2 APRGGGAAATAALDITGVHILEASVPLPEAGGNVORKGAVD-----PDKD 51  
DB 27 APMQGGGSAATPA-----ASASATPASETTVARELDGLDIQGDAPSSQPATS 76  
  
QY 52 RKEKKAAPRIT-----GWLREYSKIVCEKAKGRTTYNEVADEIYSELKSMHIGQG 106  
DB 77 KKKRGEGTRATGPDKGGRGLRQFSMKVCEKVESKGRITTYNEVADELVADELFADPNNNFAS 136  
  
QY 107 -----FDEKNIRRRVYDAFNVLIALRVIAKEKEIRWGLSNRYVEKIKKLEEV 156  
DB 137 PDPENPTPQFDEKNIRRRVYDALNVLAMDIISKDKKEIQWKLPR---TSMDSVEELK 193  
  
QY 157 KELV---NKIRNKALLQEIKEQFDLQNIKLNRQTLLESSAENVN-GIRLPFLVKTSRK 212  
DB 194 TEIIGLGRIDKKNVAYLQELQDFVGLQNLQNRNEQLYSGSNAPSGGVALPFILVQTRPH 253  
  
QY 213 ARVEIEISDSDSKFAHFENFGAPTLHDDLILSGVRNRSIGRAGR 257  
DB 254 ATVEVEISEDMQLVHFDNFSTPPELHDDSFVLK-----ALGFSGK 293

RESULT 9  
AAU72514  
ID AAU72514 standard; protein; 376 AA.

XX AAU72514;  
XX AC  
XX 26-FEB-2002 (first entry)  
XX Arabidopsis partial cell cycle protein CCP33.  
XX  
XX Cell cycle protein; CCP; cell cycle regulation; herbicide;  
KW plant growth regulator; plant development; abiotic stress; biotic stress;  
KW nutrient deprivation; pathogen attack; crop yield; immunogen.  
XX Arabidopsis thaliana.  
XX OS  
XX WO200185946-A2.  
XX PN  
XX 15-NOV-2001.  
XX PD  
XX 14-MAY-2001; 2001WO-IB001307.  
XX PF  
XX 12-MAY-2000; 2000US-0204045P.  
XX PR  
XX (CROP-) CROPDISEGN NV.  
XX PA  
XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;  
XX WPI; 2002-062249/08.  
XX DR  
XX N-PSDB; AAS96304.  
XX

PT New cell cycle protein and nucleic acid molecule encoding it useful for  
PT regulating cell cycle progression in plants and for identifying  
XX modulators which are useful as herbicides or plant growth regulators.  
XX  
XX Claim 34; Page 227-228; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the  
XX polynucleotides encoding them. CCP is useful for identifying a compound  
XX which modulates the activity of the polypeptide and which binds to the  
XX polypeptide and an anti-CCP antibody is useful for detecting the presence  
XX of CCP in a sample. A CCP modulator is useful for modulating the cell  
XX cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
XX maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP  
XX nucleic acid and polypeptide molecules are useful as modulating agents in  
XX regulating cell cycle progression in plants. CCP is useful to treat  
XX disorders characterised by insufficient or excessive production of CCP  
XX protein or production of CCP protein forms which have decreased or  
XX aberrant activity. Compounds that bind to or modulate the activity of CCP  
XX polypeptide are useful as herbicides or plant growth regulators. The  
XX polynucleotide is useful for modifying cell fate, plant development,  
XX plant morphology, biochemistry and/or physiology, the length of the G1,  
XX S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
XX stimulation or enhancement of cell division, DNA replication, seed set,  
XX seed size, seed development, tuber, fruit, leaf formation, shoot and root  
XX initiation and/or development, nodule function, dwarfism in plants,  
XX senescence, tolerance or resistance to stress. CCP, the polynucleotide  
XX and the anti-CCP antibody are useful in agriculture to modulate the  
XX protein levels or activity of a protein involved in the cell cycle due to  
XX environmental conditions, including abiotic stress such as cold, nutrient  
XX deprivation, heat, drought, salt stress, or biotic stress such as  
XX pathogen attack, to modulate e.g. enhance crop yields, and attenuate  
XX plant architecture, plant quality traits, plant reproduction and seed  
XX development, endoreduplication in storage cells, storage tissues and/or  
XX storage organs of plants or its parts. CCP is useful as an immunogen to  
XX generate antibodies. CCP protein is useful to screen for naturally  
XX occurring CCP substrates. The polynucleotide is useful for expressing CCP  
XX protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to  
XX modulate CCP activity. The present sequence represents a CCP protein of  
XX the invention

XX Sequence 376 AA;

Query Match 35.4%; Score 466.5; DB 5; Length 376;  
Best Local Similarity 46.3%; Pred. No. 5.2e-35;  
Matches 101; Conservative 38; Mismatches 64; Indels 15; Gaps 3;  
  
QY 37 GNAVORKGAVDPDKRKEKAAAPRITGWLREYSKIVCEKAKGRTTYNEVADEIYSE 96  
DB 74 GDDAGSGAGSVKVKKKRQORAGPDKTGRGLRQFSMKVCEKESKGRITTYNEVADELVAE 133  
  
QY 97 LKSMHIG-----QGFDEKNIRRRVYDAFNVLIALRVIAKEKEIRWGLSNRYVEKIKK 151  
DB 134 FALPNNDGTSPDQQQYDEKNIRRRVYDALNVLAMDIISKDKKEIQWGLRPTSLSDIEE 193  
  
QY 152 LEEVRKELVKNIRKALLQEIKEQFDLQNIKLNRQTLLESSAENVN-GIRLPFLVKT 210  
DB 194 LKNERLSLRNRIEKKTAYSQLESEQ-----RNEHLYSSGNAPS GGVALPFILVQTR 244  
  
QY 211 RKARVEIEISDSDSKFAHFENFGAPTLHDDLILSGVR 248  
DB 245 PHATVEVEISEDMQLVHFDNFSTPPELHDDNFVLKTMK 282

RESULT 10  
ADQ36855  
ID ADQ36855 standard; protein; 379 AA.  
XX ADQ36855;  
XX  
XX 07-OCT-2004 (first entry)  
XX Cell proliferation-related polypeptide #5.  
XX

KW cell proliferation related polypeptide; cell proliferation; senescence;  
XX differentiation; stress response.  
OS Oryza sativa.  
XX WO2004061122-A2.  
XX 22-JUL-2004.  
XX 23-DEC-2003; 2003WO-US041200.  
XX 26-DEC-2002; 2002US-043656P.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Cooper B;  
XX WPI; 2004-534388/51.  
XX New nucleic acid molecule encoding a cell proliferation-related  
PT polypeptide, useful for modulating cell proliferation, senescence,  
PT differentiation, development, and stress response in plants, and for  
PT producing enhanced food crops.  
XX Claim 28; SEQ ID NO 14; 408pp; English.  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a cell proliferation-related polypeptide. The nucleic acid  
CC molecule and the encoded polypeptide, and methods are useful for  
CC modulating cell proliferation, senescence, differentiation, development,  
CC and stress response in plants, and for producing enhanced food crops. The  
CC present sequence represents a cell proliferation-related polypeptide. The  
CC present sequence is published separately from the main body of the  
CC specification as EPO data.  
XX  
XX Sequence 379 AA;  
Query Match 35.4%; Score 466.5; DB 8; Length 379;  
Best Local Similarity 37.8%; Pred. No. 5.3e-35;  
Matches 111; Conservative 43; Mismatches 77; Indels 63; Gaps 8;  
QY 3 PPRGGAATAALDVTGVHLEASSVPP-----LPEAGGNVORKGAVDP----- 48  
DB 37 PPSGGAQSASTSG-----GSAGSPSRSEQHVPAAAGMAAGAAATASTPISNTFL 86  
QY 49 -----DKDRKEKAAAPRITGWLREYSKIYCEKVEAKGRITYNE 88  
DB 87 RLNDLIHGDDAPSSQAPTSKKKRGARAVGPKGGRGLRQFSMKVCEKVESGRITYNE 146  
QY 89 VADEIYSELKSMAH-----GQGFDEKNIRRVYDAFNVLIALRVIAKEKEIRW 138  
DB 147 VADELVAEFADPNNSILPPDPNPNQAQYDEKNIRRVYDALNVLNMAEIIISKDKKEIQW 206  
QY 139 MGLSNTRYEKIKLVEVRKELV---NKIRNKALLQIEKQFDLQNIKLNRQTLSESAE 195  
DB 207 KGLPR---TSINDIEDLQTELVLGKRIEKNYTLQELQDFVGMQKLIQRNEQLVSG- 262  
QY 196 NV---NGIRLPFLVKTSRKARVEIISDDSKFAHFEFNGCAPFTLHDDLILEGV 247  
DB 263 NIPSGGVALLPFLVQTRPHATVEVEISEDMQLVHFDNFSTPFELHDDSFVLKAM 316  
RESULT 11  
AAB07975  
ID AAB07975 standard; protein; 299 AA.  
XX AAB07975;  
XX 14-NOV-2000 (first entry)  
XX A corn transcription factor designated DP.  
DE DP polypeptide; transcription factor; gene transcription; cell cycle;  
XX

KW DNA replication; DRTF; differentiation-regulated transcription factor 1;  
KW transgenic plant; transformation efficiency.  
OS Zea mays.  
XX Key Location/Qualifiers  
FH Misc-difference 1 /note= "any amino acid encoded by NNN"  
FT  
XX WO2000047614-A1.  
XX 17-AUG-2000.  
XX 11-FEB-2000; 2000WO-US003651.  
XX 12-FEB-1999; 99US-0119857P.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX Lowe KS, Gordon-Kamm WJ, Bailey MA, Gregory CA;  
XX WPI; 2000-524530/47.  
XX N-PSDB; AAA59702.  
XX Transgenic plants comprising an expression cassette consisting of a DP  
PT nucleic acid, when expressed, the nucleic acid increases the efficiency  
PT of plant cell transformation by increasing the number of dividing cells  
PT in the plant.  
XX Disclosure; Page 71-72; 76pp; English.  
XX The present sequence represents a corn DP (not defined) polypeptide. The  
CC polypeptide is a transcription factor that activates transcription of  
CC numerous genes involved in DNA replication, thus playing a role in the G1  
CC to S transition in the cell cycle. Do, also called DRTF (differentiation-  
CC regulated transcription factor 1) has been shown to form specific  
CC multiprotein complexes with the retinoblastoma susceptibility protein,  
CC p107, cyclins and cdk2. The DP polynucleotide was used to produce a  
CC transgenic plant. The DP nucleic acids and proteins are useful for  
CC increasing transformation efficiency in plants by increasing the number  
CC of dividing cells in the plant (dividing cells may be more receptive to  
CC transformation)  
XX  
XX Sequence 299 AA;  
Query Match 35.0%; Score 461; DB 3; Length 299;  
Best Local Similarity 40.7%; Pred. No. 1.2e-34;  
Matches 110; Conservative 41; Mismatches 71; Indels 48; Gaps 8;  
QY 16 LDLTGVHLEASSVPLPEAGGNVORKG--AVDPDKRKEKAAAPRITGWLREYSKI 73  
DB 11 LDING-----DDAPSSQAPTSKKKRGTRAVGPKDGNR-----GLRQFSMK 51  
QY 74 VCEKVEAKGRITYNEVADEIYSELKSMAH-----IGQGFDEKNIRRVYDAFNVL 123  
DB 52 VCEKVESKGRITYNEVADELVAEFTDPNNIEAPDPNPNQAQYDEKNIRRVYDALNVL 111  
QY 124 IALRVIAKEKEIRWGLSNTRYEKIKLVEVRKELV---NKIRNKALLQIEKQFDL 180  
DB 112 MAMDIIISKDKKEIQWKLPR---TSISDTEEMKTELVLGKRIEKKSAVLAQLQDQVVL 168  
QY 181 QNLIKRN-QTLESSAENVNGIRLPFLVKTSRKARVEIISDDSKFAHFEFNGCAPFTLH 239  
DB 169 QNLIQRNEQSYSGNTSPSGVALPFLVQTRPHATVEVEISEDMQLVHFDNFSTPFELH 228  
QY 240 DLSILEGVR-----RNSIGRAGRAT 259  
DB 229 DSYVLKEMRFCCGREQHDSIQESISNGESS 258  
RESULT 12  
AAG22852  
ID AAG22852 standard; protein; 277 AA.

[illegible]



```
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.1%; Score 437; DB 3; Length 277;
Best Local Similarity 50.1%; Pred. No. 2e-32;
Matches 91; Conservative 32; Mismatches 52; Indels 6; Gaps 2;

QY 74 VCEKVFAGRTTNEVADIEYSELKSMHIG-----QGFDEKIRRRVYDAFNVLIALRV 128
DB 3 VCEKVESKRTTNEVADIEYSELKSMHIG-----QGFDEKIRRRVYDAFNVLIALRV 128
QY 129 IAKEKEIEIWMGLSNRYEKKLEBRKELVNKIRKALLQIEIKQFDDLNQIKLRNQ 188
DB 63 ISKDKKEIQWRGLPRTSLSDIEELKNERLSLRNIEKKTAYSQELSEQYVGLNQLQRNE 122
QY 189 TLESSAENVN-GIRLPFLVKTSRKARVEIISDDSKFAHFENGAFFTLHDDLSILEGV 247
DB 123 HLYSSGNAPSGGVALFPIVQTRPHATVEIVEISDMQLVHFDNFSPFELHDDNFVLKTM 182
QY 248 R 248
DB 183 K 183
```

## RESULT 13

AAU72529

ID AAU72529 standard; protein; 292 AA.

XX AAU72529;

XX 26-FEB-2002 (first entry)

DT Arabidopsis cell cycle protein CCP16.

DE

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;

KW plant growth regulator; plant development; abiotic stress; biotic stress;

KW nutrient deprivation; pathogen attack; crop yield; immunogen.

XX Arabidopsis thaliana.

XX WO200185946-A2.

XX 15-NOV-2001.

PD 14-MAY-2001; 2001WO-IB001307.

PF 12-MAY-2000; 2000US-0204045P.

XX (CROP-) CROPDISEIGN NV.

PA Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

PI WPI; 2002-062249/08.

XX N-PSDB; AAS96319.

DR New cell cycle protein and nucleic acid molecule encoding it useful for

PT regulating cell cycle progression in plants and for identifying

PT modulators which are useful as herbicides or plant growth regulators.

XX Claim 34; Fig 16; 316pp; English.

The invention relates to a novel cell cycle protein (CCP) and the polynucleotides encoding them. CCP is useful for identifying a compound which modulates the activity of the polypeptide and which binds to the polypeptide and an anti-CCP antibody is useful for detecting the presence of CCP in a sample. A CCP modulator is useful for modulating the cell cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat, maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP nucleic acid and polypeptide molecules are useful as modulating agents in regulating cell cycle progression in plants. CCP is useful to treat disorders characterised by insufficient or excessive production of CCP protein or production of CCP protein forms which have decreased or aberrant activity. Compounds that bind to or modulate the activity of CCP polypeptide are useful as herbicides or plant growth regulators. The polynucleotide is useful for modifying cell fate, plant development, plant morphology, biochemistry and/or physiology, the length of the G1, S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion, stimulation or enhancement of cell division, DNA replication, seed set, seed size, seed development, tuber, fruit, leaf formation, shoot and root initiation and/or development, nodule function, dwarfism in plants, senescence, tolerance or resistance to stress. CCP, the polynucleotide and the anti-CCP antibody are useful in agriculture to modulate the protein levels or activity of a protein involved in the cell cycle due to environmental conditions, including abiotic stress such as cold, nutrient deprivation, heat, drought, salt stress, or biotic stress such as pathogen attack, to modulate e.g. enhance crop yields, and attenuate plant architecture, plant quality traits, plant reproduction and seed development, endoreduplication in storage cells, storage tissues and/or storage organs of plants or its parts. CCP is useful as an immunogen to generate antibodies. CCP protein is useful to screen for naturally occurring CCP substrates. The polynucleotide is useful for expressing CCP protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to modulate CCP activity. The present sequence represents a CCP protein of the invention

XX

[illegible]

CC overriding the cell differentiation signals, or for altering cell shape.  
CC The present sequence represents the dimerisation partner.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

XX

**SQ Sequence 292 AA;**

Query Match	32.5%	Score 428.5;	DB 7;	Length 292;
Best Local Similarity	41.3%;	Pred. No. 1.4e-31;		
Matches 88; Conservative	49;	Mismatches 69;	Indels 7;	Gaps 2;

  

Qy	40	VORKGAVDPDKRRKKEKAAPRITGMGLRSYKIVCEKVEAKGRTTIVNEVADEIYSLKs	99
Db	27	VRRKLIVDDSEIGSEKKGGQSRTSGGGLRFQSVMCQLEAKKITTYKEVADSIISDFAT	86
Qy	100	MAHIGQ-----GFDEKNIRRRYDAPNVLIALRVIAKEKEIKRWGLSNRYEKKILE	153
Db	87	IKQNAEKPLNENYEKNIERRYYDALNVFMALDI IARDKKEIRWKGLPITCKKDVEEVK	146
Qy	154	EVRKELVNKRNKALLOETEKOFDDIQNTLKRNOTLESSAEN-VNGIRLPFVLVKTSRK	212
Db	147	MDRNKVMSVQKAAFLKELREKVSSLESLSMRNQEMVMVTQGPAGFTLPFILETINPH	206
Qy	213	ARVEIISDSKFAHPFNFGAPPTLHHDDLILSILE	245
Db	207	AVVEIISEDMLVHLDFNSTPFSVHDDAVILK	239

Search completed: February 4, 2005, 20:30:04

Job time : 168 secs

*This Page Blank (uspto)*



C;Accession: T48268  
R;Byvan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,  
submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z24490  
A;Accession: T48268  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-288 <BEV>  
A;Cross-references: UNIPROT.Q9LZ55; EMBL.ALJ62971  
A;Experimental source: cultivar Columbia; BAC clone T22P11  
C;Genetics:  
A;Map position: 5  
A;Introns: 36/2; 57/1; 71/3; 95/3; 142/3; 165/3; 198/3; 222/2  
A;Note: T22P11.60

Query Match	32.5%;	Score 428.5;	DB 2;	Length 288;
Best Local Similarity	41.3%;	Pred. No. 4.3e-21;		
Matches	88;	Conservative 49;	Mismatches 69;	Indels 7; Gaps 2;
Qy	40	VORKGAVDPDKRKKKAAAPRTGTGWLREYSKIVCEKVEAKGRTTYYNEVADEIYSLK	99	
Db	23	VRRLIIVDDSEITGSEKKGGQSRSTGGGLRQFVVMVCKLEAKKITYKEVADEIISDFAT	82	
Qy	100	MAHTGQ-----GPDEKNIRRRVYDAFNVLIALRVIAKEKKEIRMGISNRYRYKKLE	153	
Db	83	IKQNAEXPLNENYENENIRRRVYDALNVFALDIIARDKKEIRWKGLPITCKKDVSEVK	142	
Qy	154	EVRELKYNRNKKALLOETKEQFDDIQNIKLRNQTLESSAEN-VNGIRLRFVLVVKTSRK	212	
Db	143	MDRNKVNSSVQKKAFTKELREKVSSLESLSMRNQEMVMTQGPABGFTLFFILLEINPH	202	
Qy	213	ARVEIISDDSKFAHFFNGAPFTLHDDLSTLE	245	
Db	203	AVVEIIESDMOLVHLDENSTPFPSVHDDAYILK	235	

RESULT 3

T12511  
hypotheical protein DKFPz434G222.1 - human (fragment)  
N/Alternate names: E2F dimerization partner 2  
C/Species: Homo sapiens (man)  
C/Date: 23-Jul-1999 #sequence revision 23-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T12511; A57381; I37297; I39180  
R/Author(s): W.; Wlirker, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, June 1999  
A/Reference number: Z17527  
A/Accession: T12511  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-416 <ANS>  
A/Cross-references: UNIPROT:Q9UG28; EMBL:AL080206  
A/Experimental source: adult testis; clone DKFPz434G222  
R/Wu, C.L.; Zukerberg, L.R.; Ngwu, C.; Harlow, E.; Lees, J.A.  
Mol. Cell. Biol. 15, 2536-2546, 1995  
A/Title: In vivo association of E2F and DP family proteins.  
A/Reference number: A57381; MUID:95257935; PMID:7739537  
A/Accession: A57381  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 'M', 33-416 <WUH>  
A/Cross-references: GB:I40386; NID:G703084; PIDN:AAA69016.1; PID:G703085  
R/Zhang, Y.; Cheilappan, S.P.  
Oncogene 10, 2085-2093, 1995  
A/Title: Cloning and characterization of human DP2, a novel dimerization pa  
A/Reference number: I37297; MUID:95303470; PMID:7784053  
A/Accession: I37297  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA  
A/Residues: 'M', 33-143 'Q', 144-416 <ZHA1>  
A/Cross-references: EMBL:U18422; NID:G604478; PIDN:AAB60378.1; PID:G604479  
A/Accession: I39180  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA

A:Residues: 'M', 33-49 <ZHA2>  
A:Cross-references: EMBL:U35117; NID:g1008545; PIDN:AAC50642.1; PID:g1008546  
C:Genetics:  
A:Gene: GDB:TFDP2; DP2  
A:Cross-references: GDB:638814  
A:Map position: 3q23-3q23  
A:Note: DKFZp434G22.1  
C:Superfamily: transcription factor DP

	Query Match	30.1%; Score 396.5; DB 2; Length 416;
	Best Local Similarity	33.9%; Pred. No. 8.4e-19;
	Matches 101; Conservative	52; Mismatches 96; Indels 49; Gaps 9;
Qy	4 PRGAAATAAALDL-----TGWHLEASSVP	30
Dd	12 PSPGTSGATFILDLSATSRTIISTPQLTSSGSVLIGSPYTPAPAMVQTQHIAEAGTW	71
Qy	31 PLPEAGNAVORGAVDPPDKRKEKAAAARIITGWGLREYSKIYCEKVBAKGRTTYNEVA	90
Dd	72 P-----GDRCARAKFDISDFSSESKRSKKGGD-KNGKLRFHFSMKVECKQRGRGTTSYNEVA	125
Qy	91 DEIVSEL--KSMAHTG--QGDFENIRRRYYDAFNVLIALRVIAAKEKEIKWMGLSNTRYE	147
Dd	126 DELVSEFTNNHLADSAYDQKNIRRYVDALNVLMMNMIIIKEKEIKWIGLPTNSAQ	185
Qy	148 KIKLEBVRKELVNKRKALLQEETOKPDFDDLONIKLRNOTLESSAEUV----NGRLRP	203
Dd	186 ECONLETEKQRRIERIKKAQAQIELLOOTAFLNVWRNRONEQQNQGPALNSTIQLP	245
Qy	204 FVLVKTRSRKARVTIEISSDSKFPAH-FEFNGAPTPLHDLDLSILEGVRENSTIGRAGRATL	260
b	246 FTINTSRTKYDVDCSISSD-KFEVIYNFDNT-FEIHHDIENVLMRGMSFGLESKCSSL	301

## RESULT 4

E48385  
transcription factor DRTF1 chain 1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: E48585; S38372; S38371  
R:Helin, K.; Wu, C.L.; Pattaeay, A.R.; Lees, J.A.; Dynlacht, B.D.; Ngwu, C.; Harlow, E.  
Genes Dev. 7, 1850-1861, 1993  
A:Title: Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperat  
A:Reference number: E48585; MUID:94010284; PMID:8405995  
A:Accession: E48585  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trans  
A:Molecule type: mRNA  
A:Residues: 1-410 <HEL>  
A:Cross-references: UNIPROT:Q08639  
R:Girling, R.  
submitted to the EMBL Data Library, July 1993  
A:Reference number: S38372  
A:Accession: S38372  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-410 <GIR>  
A:Cross-references: EMBL:X72310; NID:g395280; PIDN:CAA51056.1; PID:g395281  
R:Girling, R.; Pattiridge, J.F.; Bandara, L.R.; Burden, N.; Totty, N.F.; Hsuan, J.J.; La T  
Nature 365, 468, 1993  
A:Title: A new component of the transcription factor DRTF1/E2F.  
A:Reference number: S38371; MUID:94019777; PMID:8413592

F:Zhang, Y.; Chellappan, S.P.  
Oncogene 10, 2085-2093, 1995  
A:Title: Cloning and characterization of human DP2, a novel dimerization partner of E2F  
A:Reference number: 137297; MUID:95303470; PMID:7784053  
A:Accession: 137297  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 'M', 33-143, 'O', 144-416 <ZHA1>  
A:Cross-references: EMBL:U18422; NID:G604478; PIDN:AAB60378.1; PID:G604479  
A:Accession: 139180  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

Query Match	28.7%;	Score 378.5;	DB 2;	Length 410;
Best Local Similarity	42.4%;	Pred. No. 1.2e-17;		
Matches	87;	Conservative	41;	Mismatches 60;
			Indels	17;
			Gaps	7



R;Barlow, K.  
submitted to the EMBL Data Library, December 1995  
A;Reference number: Z19995  
A;Accession: T25207  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-598 <WIL>  
A;Cross-references: UNIPROT:Q22703; EMBL:Z68319; PIDN:CAA92699.1; GSPDB:GN00020; CESP:T2  
A;Experimental source: clone T23G7  
C;Genetics:  
A;Gene: CESP:T23G7.1  
A;Map position: 2  
A;Introns: 3/3; 115/3; 204/2; 227/2; 260/1; 298/1; 363/1; 424/3; 478/3; 565/1

Query Match 26.8%; Score 354; DB 2; Length 598;  
Best Local Similarity 41.3%; Pred. No. 7.9e-16;  
Matches 78; Conservative 34; Mismatches 67; Indels 10; Gaps 3;

66 GLREYSKIVCEKVEAKGRITTYNVADEIYSE-----LKSMHIGQGFDEKNIRRVYDA 119  
Db GLRHFSKVKCEKVEKGLTNYNVADELVADYFQNNLIKQIDVVKQYDMKNIRRVYDA 131

120 FNVLIARLVAKKIRRMGLSNRYEKIKLEEVKELVNKRNKALLQIEKQFDD 179

132 LNVLLAMNITKSKDIRWIGLPASASQIEISRLSEKSRREASISKKQALEEVLQIVS 191

180 LQNIKLNRNTOLE---SSAENVNGIRLPVLVKTSRKARVEIISDDSKFAHFENGAFFT 236

192 YKNLVERNRNEHNGRPENDTVLHLPFLINTKEANVECSVSDKSEFLFSFD-KKFE 250

237 LHDDLSTLE 245

251 IHDDFEILK 259

RESULT 9

JC4929  
transcription factor E2F1 - human

N;Alternate names: APl2 protein; retinoblastoma-associated protein; retinoblastoma-bindin

C;Species: Homo sapiens (man)

C;Date: 22-Oct-1996 #sequence revision 01-Nov-1996 #text\_change 09-Jul-2004

C;Accession: JC4929; A45032; A42998; A42997; I54091

R;Neuman, E.; Sellers, W.R.; McNeil, J.A.; Lawrence, J.B.; Kaelin Jr., W.G.

Gene 173, 163-169, 1996

A;Title: Structure and partial genomic sequence of the human E2F1 gene.

A;Reference number: JC4929; MUID:97082961; PMID:8964493

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-437 <NEU>

A;Cross-references: UNIPROT:Q01094; GB:U47675; NID:91594281; GB:U47676; NID:91594282; GE

A;Experimental source: placenta

R;Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.

Mol. Cell. Biol. 12, 5620-5631, 1992

A;Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins

A;Reference number: A45032; MUID:93078763; PMID:1448092

A;Accession: A45032

A;Molecule type: mRNA

A;Residues: 'LCRRFRCAEDRALAEACRHGAPPPPPVTRAAARVSV', 1-312, 'S', 314-321, 'N', 323-328, 'T',

A;Cross-references: GB:S49592; NID:G260573; PIDN:AA34289.1; PID:G260574

A;Note: sequence extracted from NCBI backbone (NCBIP:119095)

A;Note: the authors are uncertain whether Met-1 is the initiator or whether translation

R;Helin, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Fattaey, A.

Cell 70, 337-350, 1992

A;Title: A cDNA encoding a PRB-binding protein with properties of the transcription fact

A;Reference number: A42998; MUID:92346720; PMID:1638634

A;Accession: A42998

A;Molecule type: mRNA

A;Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>

A;Cross-references: GB:N96577; NID:G181917; PIDN:AAA35782.1; PID:G181918

A;Experimental source: Nalm 6 pre-B leukemia cell line

A;Note: sequence extracted from NCBI backbone (NCBIN:110015, NCBIP:110016)

R;Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.

Cell 70, 351-364, 1992

A;Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-

A;Reference number: A42997; MUID:92346721; PMID:1638635

A;Accession: A42997

A;Molecule type: mRNA

A;Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>

A;Cross-references: GB:U13851; NID:9595713

A;Experimental source: Akata cells; expression vector pGEX-2TK

A;Note: sequence extracted from NCBI backbone (NCBIN:110018, NCBIP:110019)

R;Johnson, D.G.; Ohtani, K.; Nevins, J.R.

Genes Dev. 8, 1514-1525, 1994

A;Title: Autoregulatory control of E2F1 expression in response to positive and negative

A;Reference number: I54091; MUID:95047311; PMID:7958836

A;Accession: I54091

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-88, 'R', 'T', '122-123, 'TPR', '127, 'QRR', '297-299, 'PRR', '308-309, 'RA', '312, 'C', <RES>

A;Cross-references: GB:S74230; NID:9712816; PIDN:AAD14150.1; PID:G4261850

C;Genetics:

A;Gene: GDB:E2F1

A;Cross-references: GDB:134661; OMIM:189971

A;Map position: 20q11-20q11

A;Introns: 87/3 118/1 191/2 242/2 280/3 356/1

C;Keywords: DNA binding; transcription factor

F;67-108/Region: cyclin box #status predicted

F;118-190/Domain: DNA binding #status predicted <DNA>

F;191-241/Region: 7-residue repeats

Query Match 9.7%; Score 128.5; DB 2; Length 437;

Best Local Similarity 24.3%; Pred. No. 0.32;

Matches 64; Conservative 33; Mismatches 97; Indels 69; Gaps 10;

QY 2 APPRGAAAAATAALDLTGWHILEAS-----SVPLPEAGGNVQKGVDPD- 49

Db APAGGCPAPALEALLGAGALRLDLSQIVIIISAAQDASAPPAPT--GPAAPAGPCDDPL 63

QY 50 -----KDRKKKAAAPR-----ITGWLREYSKIVCEK 77

Db LIFATPQAPRPTSPAPRALGPPVKKRDLDTDHOYLAESSGPARGRHPKGVKSPG 123

QY 78 VBAGKRTTYNEVADEIYSELKSMHIGQGFDEKN-----IRRVYDAFNVLIALRVI 129

Db EKSRYETSLNLTTRF---LELLSHSDGVDLNNAAEVLKVKGRYIDITNVLGIQLI 190

QY 130 AKE-KKEIRWGLSNRYEKIKLEEVKELVNKRNKALLQIEKQFDDLQNIKLNRQ 188

Db AKKKNHIOWLQ-SHTTVGVGRLEGLTQDL-----ROLQESQQLDHLNICTTQL 231

QY 189 TLESSAENVNGIRLPVLVKTSR 211

Db RLLS--EDTDSORLAYVTCQDLR 252

RESULT 10

T31602

hypothetical protein Y48C3A.t - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T31602

R;Wallis, J.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z21046

A;Accession: T31602

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-470 <WIL>

A;Cross-references: EMBL:AL117203; NID:e1549827; PIDN:CAB55117.1; CESP:Y48C3A.t

A;Experimental source: clone Y48C3A

C;Genetics:

A;Gene: CESP:Y48C3A.t

A;Introns: 52/3; 156/2; 206/2; 237/3; 332/1; 386/1

Query Match 9.1%; Score 120.5; DB 2; Length 470;



C:Accession: J27067  
 R:McMurray, A.  
 submitted to the EMBL Data Library, January 1998  
 A:Reference number: Z20305  
 A:Accession: T27067  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-717 <WIL>  
 A:Cross-references: UNIPROT:O45959; EMBL:AL021493; PIDN:CAA16395.1; GSPDB:GNO0023; CESP:Y1A2B  
 A:Experimental source: clone Y51A2B  
 C:Genetics:  
 A:Gene: CESP:Y51A2B.6b  
 A:Map position: 5  
 A:Introns: 8/1; 33/1  
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y116A8C.33  
  

Query Match	8.5%;	Score 112.5;	DB 2;	Length 717;
Best Local Similarity	26.0%;	Pred. No. 6.6;		
Matches	57;	Conservative	35;	Mismatches 54;
				Indels 73;
				Gaps 14;

  

QY	87	NEVADBIY-SELKS-MAHIGQGFDEKIRRVYDAFNVLIALRVIAEKK-----	134
		: : :           : :	
Db	368	NSPVDKESQFYTLMLNG-----KVQDLFNVLDLSSTENFKKLKRVKMLLS	417
		: : :           : :	
QY	135	-----BIRWGLSNY-----RYEK-INKLBEVRKELVNKIRN	165
		: : :           : :	
Db	418	PINLDRHLIRVKVEGILEEQTLTSYEEKLLPFQPYRKYKLLKLLINGRKEFADKVDN	477
		: : :           : :	
QY	166	KKAL-----LOEIEKQFDLQNI-----KLRNQTLSSAENVNGIRLPFVL---VKTSRKA	214
		: : :           : :	

```

Db      418 PINLDRHLIRVKVEGILEELQQLTSLYYEKLLPFQPYRKYKGLKJLNGRKEFADKVDN 477
QY      166 KVAL---LOEIEKQFDLQNI-----KLRNQTLSSAENVNGIRLPFVL---VKTSRKAR 214

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 20:06:49 ; Search time 170 Seconds  
(without alignments)  
786.192 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAAAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1304	98.9	261	2 Q9FET1	Q9fet1 triticum sp
2	890.5	67.5	294	2 Q84VF4	Q84vf4 oryza sativ
3	838	63.5	289	2 Q8S182	Q8s182 oryza sativ
4	492	37.3	385	2 Q8LDG3	Q8ldg3 arabidopsis
5	492	37.3	385	2 Q9FNY2	Q9fny2 arabidopsis
6	476.5	36.1	353	2 Q7XZ14	Q7xz14 populus tre
7	471.5	35.7	346	2 Q84VD5	Q84vd5 oryza sativ
8	466.5	35.4	379	2 Q84VA0	Q84va0 oryza sativ
9	456	34.6	413	2 Q9LZ27	Q9lze7 arabidopsis
10	428.5	32.5	288	2 Q9LZ55	Q9lze5 arabidopsis
11	428.5	32.5	292	2 Q9FNY3	Q9fny3 arabidopsis
12	399.5	30.3	388	2 Q6R754	Q6r754 homo sapien
13	394	29.9	359	2 Q9UG28	Q9ug28 homo sapien
14	394	29.9	446	1 TDP2 HUMAN	Q14188 homo sapien
15	393	29.8	395	2 Q8BHD2	Q8bhd2 m mus muscu
16	388	29.4	385	2 Q8C537	Q8c537 mus musculu
17	382	29.0	409	2 Q632R9	Q632r9 xenopus lae
18	382	29.0	412	2 Q66K12	Q66k12 xenopus lae
19	378.5	28.7	396	2 Q9D297	Q9d297 mus musculu
20	378.5	28.7	410	1 TDP1 MOUSE	Q08639 mus musculu
21	375.5	28.5	377	1 TDP1 DROME	Q24318 drosophila
22	375.5	28.5	441	2 Q8ML56	Q8ml56 drosophila
23	375.5	28.5	445	2 Q9V6M0	Q9v6m0 drosophila
24	375.5	28.5	445	2 Q9V6M0	Q9v6m0 drosophila
25	374.5	28.4	410	1 TDP1 HUMAN	Q14186 homo sapien
26	373	28.3	409	2 Q6TNQ1	Q6tnq1 brachydanio
27	372.5	28.2	318	2 Q7XE27	Q7xe27 oryza sativ
28	372	28.2	409	2 Q803N1	Q803n1 brachydanio
29	364.5	27.6	363	2 Q44080	Q44080 drosophila
30	356	27.0	506	2 Q7PFI9	Q7pfi9 anopheles g
31	356	27.0	584	2 Q7PSH2	Q7psh2 anopheles g

32 355 26.9 310 2 Q8C8M7  
33 354 26.8 598 1 TDPI CAEEL  
34 351.5 26.6 381 2 Q9GT27  
35 349.5 26.5 290 2 Q9CYZ7  
36 345.5 26.2 386 2 Q6PBR6  
37 323 24.5 446 1 TDP2 MOUSE  
38 302 22.9 345 2 Q9NZ54  
39 234.5 17.8 657 2 Q7KWP1  
40 165 12.5 212 2 Q8RS9  
41 128.5 9.7 437 1 E2F1 HUMAN  
42 120.5 9.1 287 2 Q9BIE8  
43 120.5 9.1 412 2 Q9U289  
44 117 8.9 425 2 Q654W1  
45 116.5 8.8 864 1 RA50\_SULSO

#### ALIGNMENTS

##### RESULT 1

Q9FET1 PRELIMINARY; PRT; 261 AA.  
AC Q9FET1  
DT 01-MAR-2001 (TREMREL. 16, Created)  
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)  
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)  
DE DP protein.  
OS Triticum sp.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Triticum.  
OX NCBI\_TaxID=4569;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20362804; PubMed=11108846; DOI=10.1016/S0014-5793(00)02239-0;  
RA Ramirez-Parra E., Gutierrez C.;  
RT "Characterization of wheat DP, a heterodimerization partner of the  
RT plant E2F transcription factor which stimulates E2F2DNA binding.";  
RL FEBS Lett. 486:73-78(2000).  
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -|- SIMILARITY: Belongs to the E2F/DP family.  
DR EMBL; AJ271917; CAC19034.1; -  
DR HSSP; Q14188; 1CF7.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0005667; C:transcription factor complex; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0000074; P:regulation of cell cycle; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR009058; WING\_hlx\_DNA\_bnd.  
DR InterPro; IPR003316; E2F\_TDP.  
DR Pfam; PF02319; E2F\_TDP; 1.  
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.  
SQ SEQUENCE 261 AA; 29262 MW; 5008D8F193163A28 CRC64;

Query Match 98.9%; Score 1304; DB 2; Length 261;  
Best Local Similarity 98.9%; Pred. No. 9.5e-77;  
Matches 258; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAPPRGAAAAATAALDLTGVIHLEASSVPPIPEAGNVAQRKGVDPDKRKKKAAAP 60  
DB 1 MAPPRGAAAAATAALDLTGVIHLEASSVPPIPERGNAVQRKGVDPDKRKKKAAAP 60  
QY 61 RITGWLRLYSKIVCEKVGKRTTNEVADEIYSELKSMHITGQGFDEKTRRRYVDAF 120  
DB 61 RITGWLRLYSKIVCEKVGKRTTNEVADEIYSELKSMHITGQGFDEKTRRRYVDAF 120  
QY 121 NVLIARLVIAKEKKEIRWMGLSNRYEYKIKLEEVKELVKNIRNKALLQIEKEQFDDL 180  
DB 121 NVLIARLVIAKEKKEIRWMGLSNRYEYKIKLEEVKELVKNIRNKALLQIEKEQFDDL 180  
QY 181 QNIKLRNQTLESSAENVNGIRLPFVLVKTSRKARVEIEISDDSKFAHFENGFAPFTLHDD 240  
DB 181 QNIKLRNQTLESSAENVNGIRLPFVLVKTSRKARVEIEISDDSKFAHFENGFAPFTLHDD 240

```

QY 241 LSILEGVRNRSIGRAGRATILH 261
DB 241 LSILEGVRNRSIGRAGRATILH 261

RESULT 2
ID Q84VF4 PRELIMINARY; PRT; 294 AA.
AC Q84VF4;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE DP protein.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23111120; PubMed=14750518;
RX DOI=10.1023/B:PLAN.000007001.30865.0f;
RA Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
RA Goff S.A., Glazebrook J.
RT Identification of rice (Oryza sativa) proteins linked to the cyclin-
RT mediated regulation of the cell cycle.
RL Plant Mol. Biol. 53:273-279(2003).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AV224529; AAO72649.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q84VF4; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 294 AA; 32859 MW; D1A009564782671 CRC64;

Query Match 67.5%; Score 890.5; DB 2; Length 294;
Best Local Similarity 63.1%; Pred. No. 6.5e-50;
Matches 185; Conservative 31; Mismatches 36; Indels 41; Gaps 4;

QY 1 MAPPGGAAAAATAALDLTGCVHILEASSVPLPE----- 34
DB 1 MAPPGGAAAAAASAPGLANLLIREGAGLPSRPERPPCTSDSFAPISREGDDIPPO 60

QY 35 -----AGGNVQRK--GAVDPDKRKKKAAAPRITGWLREYSKIVCKVEAKGR 83
DB 61 KKSLSLRSVSGGGNAAREEGGA---NRNGKKEKTGAQRITGWLREYSKIVCKVEAKGR 117

QY 84 TTNEVADEIYSELKSWAHIGQGFDEKNIRRRYDAFNVLIAIRVIAKEKEIRWGLSN 143
DB 118 TTNEVADEIYFAELKSTQNGLEFDEKNIRRRYDAFNVLIAIRVIAKDKKEIKWGLTN 177

QY 144 YRYEKIKLEVRKELVNKIRNKKALQIEIKQFDLQNTKLNRQTLSSAENVNGIRLP 203
DB 178 YRYEKIQKLEVEVKELITRIKNKKLQIEIKQFDLQNTLNRQASQRPASVNGILLP 237

QY 204 FVLVKTSRKARVEIEISDDSKFAHFEFGAPFTLHDDLILLEGVRNRSIGRAG 256
DB 238 FLIIKTSRKARVEIEISDSKFAHFEFGAPFTLHDDLILLEGVRNRSIGRAG 289

RESULT 3
Q8S182 PRELIMINARY; PRT; 289 AA.
ID Q8S182
AC Q8S182;

```

```

DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Putative DP protein.
GN Name=B1144G04.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanayori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.
RT The genome sequence and structure of rice chromosome 1.
RL Nature 420:312-316(2002).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AP003335; BAB90030.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q8S182; -.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 289 AA; 32157 MW; B911BEC6E3FFDAA CRC64;

Query Match 63.5%; Score 838; DB 2; Length 289;
Best Local Similarity 60.9%; Pred. No. 1.6e-46;
Matches 179; Conservative 31; Mismatches 36; Indels 48; Gaps 6;

QY 1 MAPPGGAAAAATAALDLTGCVHILEASSVPLPE-----AGGNVQR 42
DB 1 MAPPGGAAAAAASAPGLANLLIREGAGLPSRPERGDDIPPOKKSLSVLSRGGGNAAR 60

QY 43 K--GAVDPDKRKKKAAAPRITGWLREYSKIVCKVEAKGRTTNE----- 88
DB 61 EREGA---NRNGKKEKTGAQRITGWL-----LSKKVEAKGRTTNEIMVQTSNDEVYT 111

QY 89 -----VADEIYSELKSWAHIGQGFDEKNIRRRYDAFNVLIAIRVIAKEKEIRWGLS 142
DB 112 SSGELIVADEIYFAELKSTQNGLEFDEKNIRRRYDAFNVLIAIRVIAKDKKEIKWGLT 171

QY 143 NYRYEKIKLEVRKELVNKIRNKKALQIEIKQFDLQNTKLNRQTLSSAENVNGIRL 202
DB 172 NYRYEKIQKLEVEVKELITRIKNKKLQIEIKQFDLQNTLNRQASQRPASVNGILL 231

QY 203 FVLVKTSRKARVEIEISDDSKFAHFEFGAPFTLHDDLILLEGVRNRSIGRAG 256
DB 232 PFLIIKTSRKARVEIEISDSKFAHFEFGAPFTLHDDLILLEGVRNRSIGRAG 284

RESULT 4
Q8LDG3 PRELIMINARY; PRT; 385 AA.
ID Q8LDG3
AC Q8LDG3;

```

```

[1]
RN RP MEDLINE=20562805; PubMed=11108847; DOI=10.1016/S0014-5793(00)02238-9;
RX Magyar Z., Atanassova A., de Veylder L., Rombauts S., Inze D.;
RT "Characterization of two distinct DP-related genes from Arabidopsis
RT thaliana.";
RT FBBS Lett. 486:79-87(2000).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AJ294532; CAC15484.1; -.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009059; Wing_hlx_DNA_bnd.
DR Pfam; PF023319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 385 AA; 42755 MW; 5DD84ACA04C52AF8 CRC64;

Query Match 37.3%; Score 492; DB 2; Length 385;
Best Local Similarity 47.7%; Pred. No. 5.6e-24;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQKGVADPDKRKKEKAAAPRTGTWGLREYSKIVCEKBAKGRTTYNEVADEIYSE 96
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 74 GDDAGSQGASGVKKKRGQAPGDKTGRGLRFSMKVCEKESKGRTTYNEVADELVAE 133
QY 97 LKSMAHIG-----QGFEKNIRRVYDAFNVLIALRVIAKEKKEIRWMLSNRYEKIKK 151
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 FALPNNDGTSPOQQQVDEKNIRRVYDALNVLNLMAMDIISKDKKEIQWRGLPRTSLSDIEE 193
QY 152 LEEVRELNVKIRNKALLQEIETQFDDLONIKLRNOTLESSAENVN-GIRLEFVLVTKTS 210
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 LKNERLSLRNRIEKKTAYSQELEEQVGLQNLQORNEHLYSSGNAPSGGVALPFILVQTR 253
QY 211 RKARVEIETSDSKFAHFENFGAPFTLHDDLSTILEGVR 248
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 PHATVEVEISEDQLVHFDNFSTPFELHDDNFVLKTMK 291

RESULT 6
QYXZ14 ID QTXZ14 PRELIMINARY; PRT; 353 AA.
AC QYXZ14;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Transcription factor Dp1.
GN Name=DPI;
OS Populus tremula x Populus tremuloides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
OX NCBI_TaxID=47664;
[1]
RN RP SEQUENCE FROM N.A.
RA Espinoza-Ruiz A., Saxena S., Schmidt J., Mellerowicz E., Bako L.S.,
RA Bhalerao R.P.;
RN Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RA Espinoza-Ruiz A., Bhalerao R.P.;
RN Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RA Bhalerao R.P., Sandberg G.;
RN Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AY307373; AAP73785.1; -.

```

```

DR EMBL; AF181998; AAQ13675.1; -.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription regulation.
DR SEQUENCE 353 AA; 38444 MW; 33C9F9F440D9D1F CRC64;
Query Match 36.1%; Score 476.5; DB 2; Length 353;
Best Local Similarity 42.7%; Pred. No. 5.1e-23;
Matches 112; Conservative 43; Mismatches 76; Indels 31; Gaps 6;
QY 2 APPRGAATAAALDLTGWHILEASSVPPPEAGNAVORKG--AVDPDKR 52
DB 46 SPSRSEHAVATPASDSTFURLNHLDIHADDAAT----QAAANKKKKGQRAVGADK-- 99
QY 53 KKEKAAAPRITGWLREYSKIVCEKVEAKGRITTYNEVADEIYSELKSMHI-----GQGF 107
DB 100 -----SGRLRQFSIKVCEKESKTTTYNEVADELVAEFADPSNSVSTPDQOY 149
QY 108 DEKNIRRVYDAFNVLIALRVIAKEKEIRWGLSNRYREKIKKLEVRKELVKNRK 167
DB 150 DEKNIRRVYDALNVLNLMALDIISKDKKEIQWKGLPR---TSMDSVEELK 209
QY 168 ALLOEIEKQFDLQNIKRLNOTLESSAENVN-GIRLPFLVKTSRKARVEIISDDSKFA 226
DB 210 AYLOEIEKQFVGLQNIQRNEQIYSSGNAPSGVSPFILLVQTRPHATVEVEISEDMQLV 269
QY 227 HFEFNGAPFTLHDDLSILEGVR 248
DB 270 HFDNSTPFELHDDNYVLKAMK 291
RESULT 7
Q84VD5 ID Q84VD5 PRELIMINARY; PRT; 346 AA.
AC Q84VD5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DP TF.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23111120; PubMed=14750518;
RX DOI=10.1023/B:PLAN.0000007001.30865.0f;
RA Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
RA Goff S.A., Glazebrook J.;
RA "Identification of rice (Oryza sativa) proteins linked to the cyclin-
RT mediated regulation of the cell cycle.";
RL Plant Mol. Biol. 53:273-279(2003).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AY224551; AA072671.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q84VD5; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR DNA-binding; Nuclear protein; Transcription regulation.
DR SEQUENCE 346 AA; 36907 MW; 53B09A5F68F265B0 CRC64;
Query Match 35.7%; Score 471.5; DB 2; Length 346;
Best Local Similarity 39.6%; Pred. No. 1.1e-22;
Matches 113; Conservative 44; Mismatches 81; Indels 47; Gaps 8;
QY 2 APPRGAATAAALDLTGWHILEASSVPPPEAGNAVORKGAVD-----PDKD 51
DB 27 AFGMQGGGSAATPA-----ASASASTPASTTVARLDGLDIOGDDAPSPQATS 76
QY 52 RKKEKAAAPRIT-----GWLREYSKIVCEKVEAKGRITTYNEVADEIYSELKSMHIQOG 106
DB 77 KKKKGCGTTRATGPKDGGRLRQFSIKVCEKESKTTTYNEVADELVAEFADPNNNEAS 136
QY 107 -----FDEKNIRRVYDAFNVLIALRVIAKEKEIRWGLSNRYREKIKKLESVR 156
DB 137 PDPDNPNTPFQDEKNIRRVYDALNVLNLMALDIISKDKKEIQWKGLPR---TSMDSVEELK 193
QY 157 KELV--NKIRNKKALLOEIKQFDLQNIKRLNOTLESSAENVN-GIRLPFLVKTSRK 212
DB 194 TRIIIGLKRDKKNAYLOELEDQFVGLQNLQNRNEQIYSSGNAPSGVSPFILLVQTRPH 253
QY 213 ARVEIISDDSKFAHFEFNGAPFTLHDDLSILEGVRNRSIGRAGR 257
DB 254 ATVEVEISEDMQLVHFDNSTPFELHDDSFVLK-----ALGFSGK 293
RESULT 8
Q84VAO ID Q84VAO PRELIMINARY; PRT; 379 AA.
AC Q84VAO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE E2F dimerization factor.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23111120; PubMed=14750518;
RX DOI=10.1023/B:PLAN.0000007001.30865.0f;
RA Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
RA Goff S.A., Glazebrook J.;
RA "Identification of rice (Oryza sativa) proteins linked to the cyclin-
RT mediated regulation of the cell cycle.";
RL Plant Mol. Biol. 53:273-279(2003).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AY224589; AA072709.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q84VAO; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR DNA-binding; Nuclear protein; Transcription regulation.
DR SEQUENCE 379 AA; 40893 MW; 913F06C2975FB0CB CRC64;
Query Match 35.4%; Score 466.5; DB 2; Length 379;
Best Local Similarity 37.8%; Pred. No. 2.5e-22;
Matches 111; Conservative 43; Mismatches 77; Indels 63; Gaps 8;
QY 3 PPSGGAATAAALDLTGWHILEASSVPP-----LPEAGNAVORKGAVD----- 48
DB 37 PPSGGAQASGTSG-----GSAGSPSSRSRSEQHVPAAGNAAGAAAATPISNTFL 86

```

```

QY 49 -----DKRKKEKAAAPRITGWLREYSKIVCKEYAKGRITTYNE 88
Db 87 RLNDLDIHGDAPSSQAPTSKXKRGARAVDPKGGRLGQFMSKVKESKGRITTYNE 146
QY 89 VADEIYSELKSMABI-----GGFDEKIRRRVYDAFNVLIALRVIAKEKEIRW 138
Db 147 VADELVAEFADPNNSILPPDPNPAQQYDEKIRRRVYDALNVLWAMEIISKDKKEIQW 206
QY 139 MGLSNRYEYKIKLEEVREKLV---NKIRNKALLQIEKQFDDQLQIKLRNOTLESSAE 195
Db 207 KGLPR---TSINDIEDLQTLVGLKSKRIEKNYTLQELQDFVGMQKLIQNRNQLVSG- 262
QY 196 NV--NGIRLPFVLVKTSSRKARVEISDDSKFAHFNFGAPFTLHDDLSTILEGV 247
Db 263 NIPSGGVALPFILVQTRPHATVEISEDMQLVHFDNFSTPFELHDDSFVLKAM 316

RESULT 9
Q9LZE7 PRELIMINARY; PRT; 413 AA.
AC Q9LZE7;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Transcription factor-like protein.
GN NamesF12B4_160;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidops.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AL162751; CAB83299.1; -.
DR PIR; T48364; T48364.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR InterPro; IPR003316; E2F_TDP.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 413 AA; 45604 MW; 239FDC6DAED2F723 CRC64;

Query Match 34.6%; Score 456; DB 2; Length 413;
Best Local Similarity 41.3%; Pred. No. 1.3e-21;
Matches 103; Conservative 41; Mismatches 68; Indels 34; Gaps 4;

QY 37 GNAVQKGAVDPPDKRKKEKAAAPRITGWLREYSK-----IVCKEY 79
Db 74 GDDAGSQGASGVKKKRGQRAAGPDKTGRGLRQFSKMLGISFAPIMLSKCLISICEKY 133
QY 80 AKGRTTYNEVADIYSELKSMABI-----QGFEKIRRRVYDAFNVLIALRVIAKEKK 134
Db 134 SKGRTTYNEVADLVAEAFALPNNDGTSPOQQYDEKIRRRVYDALNVLWAMDIISKDKK 193
QY 135 EIRWGLSNRYEYKIKLEEVREKLVNKNKALLQIEKQFDD-----LQNI 183
Db 194 EIQWRGLPRTSLSDI BELKNRSLRNRIEKKTAYSQELSEQYMNITDITGLSASCLNL 253
QY 184 KLRNQTLESSAENVN-GIRLPFVLVKTSSRKARVEISDDSKFAHFNFGAPFTLHDDL 242

```

```

Db 254 IQRNEHLYSSGNAPSGVALPFILVQTRPHATVEISEDMQLVHFDNFSTPFELHDDNF 313
QY 243 ILEGVR 248
Db 314 VLKTMK 319

RESULT 10
Q9LZ55 PRELIMINARY; PRT; 288 AA.
AC Q9LZ55;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE DP-2 transcription factor-like.
GN NamesF22P11_60;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidops.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AL162971; CAB85984.1; -.
DR PIR; T48268; T48268.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 288 AA; 32560 MW; EC0AD7DC06EF92B7 CRC64;

Query Match 32.5%; Score 428.5; DB 2; Length 288;
Best Local Similarity 41.3%; Pred. No. 5.2e-20;
Matches 88; Conservative 49; Mismatches 69; Indels 7; Gaps 2;

QY 40 VORKGAVDPPDKRKKEKAAAPRITGWLREYSKIVCKEYAKGRITTYNEVADIYSELKS 99
Db 23 VRKLIVDDSDSEIGSEKKGQSRTSGGLRQFSVMVQCKLEAKKITTYKEVADEIISDFAT 82
QY 100 MAHIQ-----GFDEKIRRRVYDAFNVLIALRVIAKEKKEIRWGLSNRYEYKIKLE 153
Db 83 IQQNAEKPLNENEYKIRRRVYDALNVLWAMDIIPARKKEIRWGLPITCKQDVEEVK 142
QY 154 EVRKELVNKIRNKALLQIEKQFDDQLQIKLRNOTLESSAEN-VNGIRLPFVLVKTSSRK 212
Db 143 MDRNKVMSSVQKKAAPLKEKREKVSLSMSRQEMVVKYQGPASGFTLPFILLETNPH 202
QY 213 ARVEIISDDSKFAHFNFGAPFTLHDDLSTILE 245
Db 203 AVVEIISDDSKFAHFNFGAPFTLHDDLSTILE 245

RESULT 11
Q9FNY3 PRELIMINARY; PRT; 292 AA.
AC Q9FNY3;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)

```



Qy	40	VQRKGVDPQDKRKKAAAPRITGWLGRYSKIVCEVKAAGRRTTYNSVADEIISLKLS	99
Db	27	VRRKLIIVDDDSITGESEKKGQSRTSGGGLRFFSVVWCQKLEAKKIIITKYEVADEIISDPAT	86
Qy	100	MAHTGQ-----GFDEKNTIRRRVVYDAFNVIIALRVLTAKKKETRWGGLSYRYEYKKLE	15





FT	STRAND	204	206
SQ	SEQUENCE	446.AA; 49236 MW; 19AC85BAD61DFF1 CRC64;	
	Query Match	29.9%; Score 394; DB 1; Length 446;	
	Best Local Similarity	37.5%; Pred.No.1.5e-17;	
	Matches	94; Conservative 52; Mismatches 89; Indels 16; Gaps 8	
Qy	18	LTGVHILEASSVPLPEAGNAVQRKGAVDPDKDRKEKAAPRITGWLGEYSKIYCEK	77
Dd	89	VTOHTIAEATGWVP-----GDKRKARFIDISDFSESRRSKGDK-NGKGLRHFSMKVCCK	142
Qy	78	VEAKGRTTYNEVADEIYSEL-KSMAHIG--QGFEKNIRREYYDAFNVLIALRVIAKEKK	134
Dd	143	VQRGTTSYNEVADELVSFEFTNSNNHLADSAYDOKNIRREVYPDALNVLMAMNIISKKK	202
Qy	135	EIRVMGLSNTRYEKIKLEEVKRVELVNKRINKKALLOIEIKOFDDQLONIKLRNQTLSSA	194
Dd	203	EIKWIGLPTNSAQRCQNLEIEKQRIERIKQKRAQLOELLQQIAFKNLVVORNRHQEQN	262
Qy	195	ENV-----NGRLPVLVKTRKARVETIEISDSKFAH-FEPNGAPFTLLDLSILEGVRR	249
Dd	263	QGPALNSTTQLPPIINTSRKTVIDCSISSD-KFEYLNFEDNT-FEIHDDIEVLKRMGM	320
Qy	250	NSIGRAGRATYL 260	
Dd	321	SFGLESKGCSL 331	
 RESULT 15 Q8BHD2 ID Q8BHD2 PRELIMINARY; PRT; 385 AA.			
AC	QBBDH2;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:473248C09 product:transcription factor Dp 2, full insert sequence (Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched library, clone:8430403A04 product:transcription factor Dp 2, full insert sequence).		
GN	Name=A330080J22Rik;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Lung, and Skin;		
RC	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;		
RA	Carninci P., Hayashizaki Y.;		
RA	"High-efficiency full-length cDNA cloning.";		
RL	Meth. Enzymol. 303:19-44(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Lung, and Skin;		
RC	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;		
RA	RIKEN FANTOM Consortium;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Lung, and Skin;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";		
RL	Nature 420:563-573(2002).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Lung, and Skin;		
RC	MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;		
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,		
RA	Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;		



*This Page Blank (uspto)*



```
Db 28 VTQTHIAEA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65
QY 78 VEAKGRTTYNEVADEIYSELKS-----MAHIGQGFEKNIIRRRYDAFNVLIALRVIAKEK 133
Db 66 VQKGTTSYNEVADELVSFTSNHNLAAADSQAQYDQENIRRRYDALNVLAMNIIISKEK 125
QY 134 KEIRWGLSNRYEYKIKLEEVKELVNKRNKALLQIEKQFDLQNIKLNRQTLSS 193
Db 126 KEIKWIGLPTNSAQECQNLIEKQRIERIKQKRAQLQELLQQAIAFNKLVRQNRQEQ 185
QY 194 AEN---VNG-IRLPFVLVKTSRKARVEIETSDSKFAH-FEFGAPFTLHDDLILSILEGVR 248
Db 186 NQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNDT-FEIHDDIEVLKRMG 243
QY 249 RNSIGRAGRATL 260
Db 244 MSFGLESKGCSL 255

RESULT 2
US-09-189-627A-6
; Sequence 6, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; PRIOR FILING DATE: 1998-11-10
; PRIOR FILING DATE: 1996-09-30
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 370
; TYPE: PRT
; ORGANISM: mouse
US-09-189-627A-6

Query Match 29.5%; Score 389.5; DB 3; Length 370;
Best Local Similarity 37.3%; Pred. No. 4.7e-29;
Matches 94; Conservative 51; Mismatches 74; Indels 33; Gaps 8;

QY 18 LTGVHILEASSVPPLPEAGGNAVQKGVADPDQDKRKKKAAAPRITGWLGRYSKIVCEK 77
Db 28 VTQTHIAEA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65
QY 78 VEAKGRTTYNEVADEIYSELKS-----MAHIGQGFEKNIIRRRYDAFNVLIALRVIAKEK 133
Db 66 VQKGTTSYNEVADELVSFTSNHNLAAADSQAQYDQENIRRRYDALNVLAMNIIISKEK 125
QY 134 KEIRWGLSNRYEYKIKLEEVKELVNKRNKALLQIEKQFDLQNIKLNRQTLSS 193
Db 126 KEIKWIGLPTNSAQECQNLIEKQRIERIKQKRAQLQELLQQAIAFNKLVRQNRQEQ 185
QY 194 AEN---VNG-IRLPFVLVKTSRKARVEIETSDSKFAH-FEFGAPFTLHDDLILSILEGVR 248
Db 186 NQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNDT-FEIHDDIEVLKRMG 243
QY 249 RNSIGRAGRATL 260
Db 244 MSFGLESKGCSL 255

RESULT 4
US-08-723-415B-8
; Sequence 8, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: LaThangue, Nicholas B.
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
```



```
/ TITLE OF INVENTION: THEREOF
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHYE P.C.
/ STREET: 1100 No. 5859199th Glebe Rd. 8th floor
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22201-4741
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/723,415B
/ FILING DATE: 30-SEP-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9610195.1
/ FILING DATE: 15-MAY-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crawford, Arthur R.
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 117-220
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-816-4000
/ TELEFAX: 703-816-4100
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 446 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-723-415B-2

Query Match 29.5%; Score 389; DB 2; Length 446;
Best Local Similarity 38.6%; Pred. No. 6.9e-29;
Matches 98; Conservative 48; Mismatches 86; Indels 22; Gaps 10;

QY 18 LTGVHILEASSVPLPEAGNAVQRKGA---VDPDKDRKKEKAAAPRITGWLREYSKIV 74
Db 89 VTQTHIAEA-----AGWVPSRKRAREFIDSFSESKRSKKGDK-NGKGLRHFSMKV 139

QY 75 CEKVEAKGRTTYNEVADEIYSEL-KSAHIG--QGFDEKNIRRVYDAFNVLIARVIAK 131
Db 140 CEKVQRKGTTSYNEVADELVSFTNSNNHLAADSAYDQENIRRVYDALNVLMMNIISK 199

QY 132 EKKEIRMMGLSNRYEKKLEEVKRLVKNIRKALLQIEKQFDDLQNIKLNRQTLE 191
Db 200 EKKEIKWIGLPTNSAQEQCNLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNE 259

QY 192 SSAEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH--FEFNGAPFTLHDDLILSG 246
Db 260 QQOQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKR 317

QY 247 VRNSIGRAGRATL 260
Db 318 MGMSFGLSGKCSL 331

US-08-723-415B-2

RESULT 8
US-09-189-627A-2
; Sequence 2, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: mouse
; US-09-189-627A-2

Query Match 29.5%; Score 389; DB 3; Length 446;
Best Local Similarity 38.6%; Pred. No. 6.9e-29;
Matches 98; Conservative 48; Mismatches 86; Indels 22; Gaps 10;

QY 18 LTGVHILEASSVPLPEAGNAVQRKGA---VDPDKDRKKEKAAAPRITGWLREYSKIV 74
Db 89 VTQTHIAEA-----AGWVPSRKRAREFIDSFSESKRSKKGDK-NGKGLRHFSMKV 139

QY 75 CEKVEAKGRTTYNEVADEIYSEL-KSAHIG--QGFDEKNIRRVYDAFNVLIARVIAK 131
Db 140 CEKVQRKGTTSYNEVADELVSFTNSNNHLAADSAYDQENIRRVYDALNVLMMNIISK 199

QY 132 EKKEIRMMGLSNRYEKKLEEVKRLVKNIRKALLQIEKQFDDLQNIKLNRQTLE 191
Db 200 EKKEIKWIGLPTNSAQEQCNLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNE 259

QY 192 SSAEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH--FEFNGAPFTLHDDLILSG 246
Db 260 QQOQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKR 317

QY 247 VRNSIGRAGRATL 260
Db 318 MGMSFGLSGKCSL 331

US-09-189-627A-2

RESULT 9
US-09-710-861-2
; Sequence 2, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: mouse
; US-09-710-861-2

Query Match 29.5%; Score 389; DB 3; Length 446;
Best Local Similarity 38.6%; Pred. No. 6.9e-29;
Matches 98; Conservative 48; Mismatches 86; Indels 22; Gaps 10;

QY 18 LTGVHILEASSVPLPEAGNAVQRKGA---VDPDKDRKKEKAAAPRITGWLREYSKIV 74
Db 89 VTQTHIAEA-----AGWVPSRKRAREFIDSFSESKRSKKGDK-NGKGLRHFSMKV 139

QY 75 CEKVEAKGRTTYNEVADEIYSEL-KSAHIG--QGFDEKNIRRVYDAFNVLIARVIAK 131
Db 140 CEKVQRKGTTSYNEVADELVSFTNSNNHLAADSAYDQENIRRVYDALNVLMMNIISK 199

QY 132 EKKEIRMMGLSNRYEKKLEEVKRLVKNIRKALLQIEKQFDDLQNIKLNRQTLE 191
Db 200 EKKEIKWIGLPTNSAQEQCNLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNE 259

QY 192 SSAEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH--FEFNGAPFTLHDDLILSG 246
Db 260 QQOQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKR 317

QY 247 VRNSIGRAGRATL 260
Db 318 MGMSFGLSGKCSL 331

US-09-710-861-2
```



QY 132 EKKEIRWGLSNRYEYKIKLEVRKLVNKNRKKALLQEIIEKQFDDQLQNIKLRNQTLE 191  
Db 200 EKKEIKWGLPTNSAQEONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNE 259  
QY 192 SSAEN---VNG-IRLPLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEG 246  
Db 260 QOQGGPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFONT-FEIHDDIEVLKR 317  
QY 247 VRNSIGRAGRATL 260  
Db 318 MGMSFGLESKCSL 331

RESULT 10  
US-09-949-016-9220  
; Sequence 9220, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9220  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9220

Query Match 29.2%; Score 385; DB 4; Length 331;  
Best Local Similarity 38.7%; Pred. No. 1.1e-28;  
Matches 92; Conservative 48; Mismatches 84; Indels 14; Gaps 8;

QY 34 EAGGNVQRKGA---VDPDKRKKKAAAPRITGWLREYSKIVCEKVKAKGRTTYNEVA 90  
Db 67 ESQSNKSDRKARKFIDSDFSKSKKGDGK-NGKGLRHFMSKVCVKQKRTTTSNEVA 125  
QY 91 DEIYSEL-KSMAHIG--QGDEKNIRRVYDAFNVLIALRVIAKEKEIRWGLSNRYE 147  
Db 126 DELVSEFTNSNNHLAADSAYDQKNIRRVYDALNVLMMNIIISKEKKEIKWGLPTNSAQ 185  
QY 148 KIKKLEVRKLVNKNRKKALLQEIIEKQFDDQLQNIKLRNQTLESSAENV---NGIRLP 203  
Db 186 EONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNEQOQGGPPALNSTIQLP 245  
QY 204 FVLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEGVRNSIGRAGRATL 260  
Db 246 FIIINTSRKTVIDCSISSD-KFEYLFNFONT-FEIHDDIEVLKMGMSFGLESKCSL 301

RESULT 11  
US-08-723-415B-4  
; Sequence 4, Application US/08723415B  
; Patent No. 5859199  
; GENERAL INFORMATION:  
; APPLICANT: Lathangue, Nicholas B.  
; APPLICANT: delaluna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
; THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5859199th Giebe Rd. 8th floor  
; CITY: Arlington

STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-723-415B-4

Query Match 29.1%; Score 384; DB 2; Length 369;  
Best Local Similarity 37.5%; Pred. No. 1.6e-28;  
Matches 94; Conservative 50; Mismatches 75; Indels 32; Gaps 9;

QY 18 LTCVHILEASSVPPPLPEAGNAVQRKGVDPDKRKKKAAAPRITGWLREYSKIVCEK 77  
Db 28 VTQTHIAEAA-----GWVPSKSKKGDGK-----NGKGLRHFMSKVCVK 65  
QY 78 VEAKGRTTYNEVADEIYSEL-KSMAHIG--QGDEKNIRRVYDAFNVLIALRVIAKEK 134  
Db 66 VORKGTTSTYNEVADELVSFTNSNNHLAADSAYDQKNIRRVYDALNVLMMNIIISKEK 125  
QY 135 EIRWGLSNRYEYKIKKLEVRKLVNKNRKKALLQEIIEKQFDDQLQNIKLRNQTLESSA 194  
Db 126 EIKWGLPTNSAQEONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNRQNEQON 185  
QY 195 EN---VNG-IRLPLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEGVR 249  
Db 186 QGPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFONT-FEIHDDIEVLKMG 243  
QY 250 NSIGRAGRATL 260  
Db 244 SFGLESKCSL 254

RESULT 12  
US-09-189-627A-4  
; Sequence 4, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189,627A  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4

```

; LENGTH: 369
; TYPE: PRT
; ORGANISM: mouse
US-09-189-627A-4

Query Match          29.1%; Score 384; DB 3; Length 369;
Best Local Similarity 37.5%; Pred. No. 1.6e-28;
Matches 94; Conservative 50; Mismatches 75; Indels 32; Gaps 9;

QY 18 LTGVHILEASSVPPLPEAGGNVQRKGAVDPKDRKKEKAAAPRITGWLREYSKIVCEK 77
Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65

QY 78 VEAKGRTTYNEVADIEYSEL-KSMAHIG--QGFEKNIRRVYDAFNVLIALRVIAKEKK 134
Db 66 VQRKGTTSYNEVADIEVSEFTSNNNHLAADSAYDQENIRRVYDALNVLMMANNIISKEKK 125

QY 135 EIRWMGLSNRYEYKIKLEEVKELVNKRKALLQEIKEQFDDLQNIKLNRNOTLESSA 194
Db 126 EIKWIGLPTNSAQEQCNLEIEKQRIERIKQKRAQLQELLQQIAFAKNLVQRNRQEQON 185

QY 195 EN---VNG-IRLPVLVKTARKARVEIEISDDSKFAH-PEFNGAPFTLHDDLSTILEGVR 249
Db 186 QGPPAVNNTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFDNT-FSIHDDIEVLKRMGM 243

QY 250 NSIGRAGRATL 260
Db 244 SFGLESGKCSL 254

RESULT 13
US-09-710-861-4
; Sequence 4, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 4
; LENGTH: 369
; TYPE: PRT
; ORGANISM: mouse
US-09-710-861-4

Query Match          29.1%; Score 384; DB 3; Length 369;
Best Local Similarity 37.5%; Pred. No. 1.6e-28;
Matches 94; Conservative 50; Mismatches 75; Indels 32; Gaps 9;

QY 18 LTGVHILEASSVPPLPEAGGNVQRKGAVDPKDRKKEKAAAPRITGWLREYSKIVCEK 77
Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65

QY 78 VEAKGRTTYNEVADIEYSEL-KSMAHIG--QGFEKNIRRVYDAFNVLIALRVIAKEKK 134
Db 66 VQRKGTTSYNEVADIEVSEFTSNNNHLAADSAYDQENIRRVYDALNVLMMANNIISKEKK 125

QY 135 EIRWMGLSNRYEYKIKLEEVKELVNKRKALLQEIKEQFDDLQNIKLNRNOTLESSA 194
Db 126 EIKWIGLPTNSAQEQCNLEIEKQRIERIKQKRAQLQELLQQIAFAKNLVQRNRQEQON 185

QY 195 EN---VNG-IRLPVLVKTARKARVEIEISDDSKFAH-PEFNGAPFTLHDDLSTILEGVR 249

```

RESULT 15  
US-08-428-131-2  
; Sequence 2, Application US/08428131  
; Patent No. 5863757  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 410 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-131-2

Query Match 28.7%; Score 378.5; DB 2; Length 410;  
Best Local Similarity 42.4%; Pred. No. 6.2e-28;  
Matches 87; Conservative 41; Mismatches 60; Indels 17; Gaps 7;  
QY 50 KKKKKKAAAPRTITGWLREYSKIVEKERTTYNEVADEIYSELKSM-AHI---GQ 105  
Db 105 RRRKGEK-----NGKLRHFSMKVCKVKQKRTTTSYNEVADELVAEFGAADNHILPNES 158  
QY 106 GFDEKNIRRVYDAFNVLIALRVIAKEKEIRWMGLSNRYEYKIKKLEEVYRKELVKNIRN 165  
Db 159 AYDQKNIRRVYDALNVLAMNIIKKEKEIKWIGLPTNSAQECQNEVERQRRLEIKQ 218  
QY 166 KKALQELQEKQFDDQNIKRNQTLLESSAENV----NGRLPVLVKTSRKARVEIISD 221  
Db 219 KOSQLQELIQQIAFNKLVORNRQAEQARRPPPPNSVHLPIIVNTSRKTVIDCSIN 278  
QY 222 DSKFAH-FEFNGAPFTLHDDLSILE 245  
Db 279 D-KFEYLFNFNT-FEIHDDIEVK 301

Search completed: February 4, 2005, 20:34:32  
Job time : 44 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 4, 2005, 20:33:05 ; Search time 132 Seconds  
(without alignments)  
644.003 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGGAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	63.8	369	16	US-10-437-963-136371
2	464	35.2	336	15	US-10-425-114-46555
3	453.5	34.4	320	15	US-10-424-599-186648
4	438	33.2	314	15	US-10-424-599-185947
5	374.5	28.4	355	14	US-10-106-698-4846
6	374	28.4	320	16	US-10-767-701-49206
7	372.5	28.2	318	16	US-10-437-963-166158
8	361	27.4	222	15	US-10-425-114-36974
9	354	26.8	575	9	US-09-220-091-7
10	342.5	26.0	207	15	US-10-425-114-71403
11	337.5	25.6	263	16	US-10-437-963-167076
12	299.5	22.7	405	14	US-10-053-248-24
13	299.5	22.7	405	16	US-10-345-837-24

14	225.5	17.1	165	15	US-10-424-599-234773	Sequence 234773,
15	203.5	15.4	74	14	US-10-214-188-10	Sequence 10, Appl
16	150	11.4	134	16	US-10-437-963-166159	Sequence 166159
17	133	10.1	37	9	US-09-900-147-1	Sequence 1, Appli
18	113.5	8.6	250	15	US-10-282-122A-46532	Sequence 46532, A
19	112.5	8.5	1156	15	US-10-369-493-43	Sequence 43, Appl
20	111.5	8.5	1178	15	US-10-282-122A-52434	Sequence 52434, A
21	111	8.4	444	15	US-10-389-566-435	Sequence 435, App
22	110.5	8.4	123	15	US-10-424-599-257840	Sequence 257840,
23	110	8.3	270	15	US-10-424-599-248849	Sequence 248849,
24	108.5	8.2	1957	15	US-10-369-493-2070	Sequence 2070, Ap
25	107	8.1	718	16	US-10-437-963-196494	Sequence 196494,
26	106	8.0	397	15	US-10-389-566-691	Sequence 691, App
27	105	8.0	317	15	US-10-389-566-692	Sequence 692, App
28	104.5	7.9	681	16	US-10-767-701-43781	Sequence 43781, A
29	104.5	7.9	965	10	US-09-842-484A-2	Sequence 2, Appli
30	104.5	7.9	965	10	US-09-842-484A-4	Sequence 4, Appli
31	104.5	7.9	965	14	US-10-184-485-3	Sequence 3, Appli
32	104.5	7.9	965	14	US-10-217-613-3	Sequence 3, Appli
33	104.5	7.9	965	14	US-10-217-613-8	Sequence 8, Appli
34	104.5	7.9	965	16	US-10-642-248-4	Sequence 4, Appli
35	103.5	7.8	324	9	US-09-220-091-9	Sequence 9, Appli
36	103	7.8	450	15	US-10-369-493-13626	Sequence 13626, A
37	102.5	7.8	563	15	US-10-425-114-71965	Sequence 71965, A
38	102.5	7.8	1069	15	US-10-424-599-259331	Sequence 259331,
39	101.5	7.7	972	9	US-09-879-959-10	Sequence 10, Appl
40	101.5	7.7	972	14	US-10-011-768B-9	Sequence 9, Appli
41	101.5	7.7	972	14	US-10-011-771B-9	Sequence 9, Appli
42	101.5	7.7	972	14	US-10-172-527-10	Sequence 10, Appl
43	101.5	7.7	972	14	US-10-326-185-95	Sequence 95, Appl
44	101.5	7.7	972	15	US-10-309-560-8	Sequence 8, Appli
45	101.5	7.7	972	16	US-10-642-248-2	Sequence 2, Appli

## ALIGNMENTS

### RESULT 1

US-10-437-963-136371  
; Sequence 136371, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 136371  
; LENGTH: 369  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_37957C.1.pep  
US-10-437-963-136371

Query Match 63.8%; Score 842; DB 16; Length 369;  
Best Local Similarity 50.4%; Pred. No. 2e-66;  
Matches 185; Conservative 32; Mismatches 36; Indels 114; Gaps 5;

QY 1 MAPPRGGAATAALDLTG.....SILEGVRNSIGRAGRATLH-----34

DB 1 MAPPGDAAAAAAGPLANLIREGAGLPSRPPFPCTSDSFAPISREGDDIPQ 60

QY 35 -----AGGNAVQRK--GAVDPDKRKKEKAAAPRITWGGLREYSKI-----73

```
Db 61 KKSVSLSRSGGNAAREEGGA---NRCKKXTGAORITGWLGRFSKIGTKLPDGL 117
QY 74 -----VCEKVEAKGRTTYN 87
Db 118 SRPGKHSYCAEVAVFTLTPYGGKPGNGRRAVTGGWVPGVDFCVSKKVEAKGRTTYN 177
QY 88 E-----VADEIYSELKSMHAHIGGPFDEKIRRRVYDAFNVLIALRVI 129
Db 178 EIMVQTSNDVYTSSELIVADEIFAELKSITQNGLEFDEKIRRRVYDAFNVLIALRVI 237
QY 130 AKEKKEIRWGLSNRYEYKLEEVYKELVNRKKNKALLOBIEKQFDLQNIKLNRQT 189
Db 238 AKDKKEIKWGLNRYEYKLEEVYKELVNRKKNKALLOBIEKQFDLQNIKLNRQA 297
QY 190 LESSAENVGIRLPVLVKTSKARVEIISDDSKFAHFEFGAPFTLHDDLILSILEGVR 249
Db 298 SQRPASVNGILLPFLINKTSRKARVEIBISEDSKFAHFEFGAPFTMDDVSILEAR 357
QY 250 NSIGRAG 256
Db 358 NNKGRAG 364
```

## RESULT 2

```
US-10-425-114-46555
; Sequence 46555, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46555
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700347688_FLI.pap
US-10-425-114-46555
```

```
Query Match 35.2%; Score 464; DB 15; Length 336;
Best Local Similarity 39.9%; Pred. No. 8e-33;
Matches 107; Conservative 45; Mismatches 76; Indels 40; Gaps 6;

QY 5 RGGAAAAATAALDLTGWHILEASSVPLPEAGGNVQKGAVDPKD-----RKK 54
Db 22 RGGPAPSASVS-----TPASESTVARRLNGLDPHGDDAPSSQPATSKKK 67
QY 55 E---KAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHAH----- 103
Db 68 KRGARAVGPDKNRGLRQFSMRVCEKVEKGRRTTYNEVADELVAEF-SDPNINIDSPDPD 126
QY 104 ---GGGFDEKNTRRVYDAFNVLIALRVIAKEKEIRWGLSNRYEYKLEEVYKELV 160
Db 127 NSQAQYDEKNTRRVYDAFNVLIALRVIAKEKEIRWGLSNRYEYKLEEVYKELV 186
QY 161 NKIRNKALIQEIKQFDLQNIKLNRQTLESSAENVN-GIRLPVLVKTSKARVEIEI 219
Db 187 GRIDKKNAYLDQDQYACIQLNVRNEQLYSGDAPSGGVALPFLVQTRPHATVEVEI 246
QY 220 SDDSKFAHFEFGAPFTLHDDLILSILEGV 247
Db 247 SEDMQVHFDFNSTPTPELQDQSFVLKAM 274
```

## RESULT 3

```
US-10-424-599-186648
; Sequence 186648, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 186648
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(320)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139556C.1.pap
US-10-424-599-186648
```

```
Query Match 34.4%; Score 453.5; DB 15; Length 320;
Best Local Similarity 42.6%; Pred. No. 6.5e-32;
Matches 104; Conservative 42; Mismatches 73; Indels 25; Gaps 5;

QY 27 SSVPP-----LPEAGGNVQKGAVDPKDKRKEKAAAPRITGWLREYS 71
Db 21 NSLPPVPVDTILKLNHLVDHADDAGSHASLAGS---KKKKGQRAVGDPDKSGRGLRQFS 77
QY 72 KIVCEKVEAKGRTTYNEVADEIYSELKSMHAH-----GGGFDEKNTRRVYDAFNVLIAL 126
Db 78 MKVCEKVESGRRTTYNEVADELVAEFSEPSNSLPPDQOYDEKNTRRVYDALNVLAM 137
QY 127 RVIAKEKEIRWGLSNRYEYKLEEVYKELVNRKKNKALLOBIE-KQFDDLQNIKL 185
Db 138 DIISKDKKEIQWRGLPRTTVNDIELKTERLGLNRNRIEKKTAYLOLEXEQFTGLQNLIQ 197
QY 186 RNQTLSSAENVNG-IRLPVLVKTSKARVEIEISDDSKFAHFEFGAPFTLHDDLIL 244
Db 198 RNEQLYSSRNPPSGSVSLPFIQVQTRPHATGGMGEISEDMQVHFDFNSTPTPELHDDNYVL 257
QY 245 EGVR 248
Db 258 KAMK 261
```

## RESULT 4

```
US-10-424-599-185947
; Sequence 185947, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 185947
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138923C.1.pep
US-10-424-599-185947

Query Match      33.2%; Score 438; DB 15; Length 314;
Best Local Similarity 45.9%; Pred. No. 1.5e-30;
Matches 94; Conservative 39; Mismatches 66; Indels 6; Gaps 2;

QY 50 KDRKKEKAAAPRITGWLREYSKIVCEKAEKGRTTYNEVADEIYSELKSMH-----IG 104
DB 54 KRRGGQAGGPKDSGRCLAYFVKCEKVESRGRSYNEVADELVAEFADPINGVSTPDQ 113

QY 105 QPDEKNIIRRVYDAFNLIALRVIAKEKEIRWGLSNRYRYEKIKLBEVRKELVNKIR 164
DB 114 QQYDEKNIIRRVYDALNVLNLMAMDIIISKKEIQWGLPRTSLSDIBELKSERLGLNRIE 173

QY 165 NKALLQEIETKQFDDLLQNIKLNRQTLSSAENVN-GIRLPFLVKTSRKARVEIISDDS 223
DB 174 KKAAYLQELBEEQYIGLQKIQORNEQYSSGNAPNGVSLPFIIVQTRPHATVEVEISED 233

QY 224 KFAHFEFGAPFTLHDDLSTILEGVR 248
DB 234 QLVHFDYICAPFEMHDDNVVLKAMK 258

RESULT 5
US-10-106-698-4846
; Sequence 4846, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 4846
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (342)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (348)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (351)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (352)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4846

Query Match      28.4%; Score 374.5; DB 14; Length 355;
Best Local Similarity 42.0%; Pred. No. 8.1e-25;
Matches 86; Conservative 42; Mismatches 60; Indels 17; Gaps 7;

QY 50 KDRKKEKAAAPRITGWLREYSKIVCEKAEKGRTTYNEVADEIYSELKSM-AHI---GQ 105
DB 111 RNRKGEK-----NGKGLRHFSKVKCEKQKQRTTSYNEVADELVAEFSAADNHILPNES 164

QY 106 GFDEKNIIRRVYDAFNLIALRVIAKEKEIRWGLSNRYRYEKIKLBEVRKELVNKIR 165
DB 165 AYDQKNIIRRVYDALNVLNLMAMNIIISKEKEIKWIGLPTNSAQRCQNLEVERQRRLRIKQ 224
```

```
QY 166 KKALLQEIETKQFDDLLQNIKLNRQTLSSAENV-----NGIRLPFLVKTSRKARVEIISD 221
DB 225 KQSQLELILQIIAFKLVQRNHRHAEQASRPPPPNSVIHLPIIIVNTSKTKTVIDCSISN 284

QY 222 DSKFAH-FEFGAPFTLHDDLSTILE 245
DB 285 D-KFEYLFNFNDNT-FEIHDDIEVLK 307

RESULT 6
US-10-767-701-49206
; Sequence 49206, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 49206
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-039-P1-K1-D2.pep
US-10-767-701-49206

Query Match      28.4%; Score 374; DB 16; Length 120;
Best Local Similarity 67.5%; Pred. No. 2e-25;
Matches 77; Conservative 17; Mismatches 18; Indels 2; Gaps 2;

QY 138 WMGLSNRYRYEKIKLBEVRKELVNKIRNKALLQEIETKQFDDI-QNITKLNRQTLSSAENV 197
DB 2 WACL-HFOYEYI-KLEETRKMIRVKNKKLIQEIETKQFDDI-QNITKLNRQTLSSAENV 59

QY 198 NGIRLPFLVKTSRKARVEIISDDSKEAHPFENGAPFTLHDDLSTILEGVRNS 251
DB 60 NGICLPFLLVKASRKARVEIISENSKPAQFDPNCTPFTLHDDVSILEAIRCNN 113

RESULT 7
US-10-437-963-166158
; Sequence 166158, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166158
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64895C.1.pep
US-10-437-963-166158

Query Match      28.2%; Score 372.5; DB 16; Length 318;
Best Local Similarity 37.7%; Pred. No. 1e-24;
```

RESULT 9  
US-03-220-091-7  
; Sequence 7, Application US/09220091  
; Patent No. US20020064523A1  
; GENERAL INFORMATION:  
; APPLICANT: H. Robert Horvitz  
; APPLICANT: Craig Ceol  
; APPLICANT: Xiaowei Lu

	Query Match	26.0%	Score 342.5	DB 15	Length 207
	Best Local Similarity	44.4%	Pred. No. 2.7e-22		
	Matches	75	Conservative 29	Mismatches 48	Indels 17
			Gaps 4		
Qy	105	QGFDEKNIRRVYDAFNVLIALRVIAKKEIRWGLSNRYEKIKLEEVKELV---	N 161		
		:                   :   :			
Db	1	QQYDEKNIRRVYDALNVLAMDIISKDKKEIQWKLPR---	TSISDIEEMKTELVLGKG 57		
Qy	162	KIRNKKALLOBIEQFPDDLQNIKLNR--QTLESSAENVNGIRLPFLVKTSTRKARVEIET	220		
		:                   :   :			
Db	58	RIEKKSAYLOBQOYVGLQNLIOQRNBSQYSGNTPSGVALPFLVQTRPRTTVEVEIS	117		
		:                   :   :			



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-24

Query Match      22.7%; Score 299.5; DB 14; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCLRLSMKMWETVQKGTTCQEVVGVGLVAKFRAASNHASPNESA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAMNIIISREKKIKWIGLTNTSAQNQNCLRVERQKRLERIKQK 214
QY 167 KALLQIEKQFDDQLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRNQYVERQVSRPLPNSVIHVFPFIISSSKKTVCINCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 13
US-10-345-837-24
; Sequence 24, Application US/10345837
; Publication No. US20040137440A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 5589
; CURRENT APPLICATION NUMBER: US/10/345,837
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 10/053,248
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-345-837-24

Query Match      22.7%; Score 299.5; DB 16; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCLRLSMKMWETVQKGTTCQEVVGVGLVAKFRAASNHASPNESA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAMNIIISREKKIKWIGLTNTSAQNQNCLRVERQKRLERIKQK 214
QY 167 KALLQIEKQFDDQLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRNQYVERQVSRPLPNSVIHVFPFIISSSKKTVCINCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 14
US-10-424-599-234773
; Sequence 24, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-24

Query Match      22.7%; Score 299.5; DB 14; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCLRLSMKMWETVQKGTTCQEVVGVGLVAKFRAASNHASPNESA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAMNIIISREKKIKWIGLTNTSAQNQNCLRVERQKRLERIKQK 214
QY 167 KALLQIEKQFDDQLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRNQYVERQVSRPLPNSVIHVFPFIISSSKKTVCINCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 13
US-10-345-837-24
; Sequence 24, Application US/10345837
; Publication No. US20040137440A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 5589
; CURRENT APPLICATION NUMBER: US/10/345,837
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 10/053,248
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-345-837-24

Query Match      22.7%; Score 299.5; DB 16; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCLRLSMKMWETVQKGTTCQEVVGVGLVAKFRAASNHASPNESA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAMNIIISREKKIKWIGLTNTSAQNQNCLRVERQKRLERIKQK 214
QY 167 KALLQIEKQFDDQLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRNQYVERQVSRPLPNSVIHVFPFIISSSKKTVCINCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 14
US-10-424-599-234773
; Sequence 24, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
NUMBER OF SEQ ID NOS: 285684  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 234773  
LENGTH: 165  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_54029C.1.pap  
US-10-424-599-234773

Query Match 17.1%; Score 225.5; DB 15; Length 165;  
Best Local Similarity 39.9%; Pred. No. 5.2e-12; Indels 23; Gaps 3;  
Matches 55; Conservative 20; Mismatches 40;

QY 1 MAPPGGMAAAATAALDTGWHILEASSVPPLPEAGGNAVORKGAVDPDKDRKKKAAAP 60  
DB 46 MATPADSTFLRLNHLIDHG-----DDAGSQGTVAK-----KKKRGQRAVGS 87

QY 61 RITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMAH-----IGGDFDEKNIRRR 115  
DB 88 DKSGRGLRQPSMKVCEKESKGRRTTYNEVADELVAEFADPDPINGVSTPDQOQYDEKNIRRR 147

QY 116 VYDAFNVLIARVIAKEK 133  
DB 148 VYDALNVLMAMDIIISKDK 165

RESULT 15  
US-10-214-188-10  
Sequence 10, Application US/10214188  
Publication No. US20030022260A1  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
HJMANNS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids

TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-214-188-10

Query Match 15.4%; Score 203.5; DB 14; Length 74;  
Best Local Similarity 58.3%; Pred. No. 1.6e-10; Indels 5; Gaps 3;  
Matches 42; Conservative 13; Mismatches 12;

QY 74 VCEKVEAKGRTTYNEVADEIYSELKSM-AHI---GQGFDEKNIRRRVYDAFNVLIARV 129  
DB 4 VCEK-ORKGTTSYNEVADELVAEFSADNHILPNESAYDQKNIRRRVYDALNVLMAMNII 62

QY 130 AKEKKEIRWMGL 141  
DB 63 SKEKKEIKWIGL 74

Search completed: February 4, 2005, 20:45:29  
Job time : 134 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 6, 2005, 20:16:37 ; Search time 4657 Seconds  
(without alignments)  
2715.656 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGGAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgn2\_1/USFIO.spool/US1008830/runat\_04022005.122602.29559/app\_query.fasta.1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1008830 @CGN 1.1 3731 @runat 04022005.122602.29559 -NCPU=3  
-NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.ste.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1314	99.6	1089	6 AX100704	AX100704 Sequence
2	1304	98.9	1083	8 TSP271917	AJ271917 Triticum
3	890.5	67.5	885	8 AY224529	AY224529 Oryza sat
4	886.5	67.2	1174	8 AK111611	AK111611 Oryza sat

#### ALIGNMENTS

RESULT 1	AX100704	Sequence 1 from Patent WO0121644.	1089 bp	DNA	linear	PAT 10-APR-2001
LOCUS	AX100704	Sequence 1 from Patent WO0121644.				
DEFINITION	AX100704	Sequence 1 from Patent WO0121644.				
ACCESSION	AX100704	Sequence 1 from Patent WO0121644.				
VERSION	AX100704.1	GI:13619652				
KEYWORDS						
SOURCE						
ORGANISM						
Triticum monococcum						
Triticum monococcum						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;						
Poideae; Triticeae; Triticum.						
REFERENCE						
1						
AUTHORS						
Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.						
TITLE						
Wheat dp proteins and uses thereof						
JOURNAL						
Patent: WO 0121644-A 1 29-MAR-2001;						
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)						
FEATURES						
Location/Qualifiers						
1..1089						
/organism="Triticum monococcum"						
/mol_type="unassigned DNA"						
/db_xref="taxon:4568"						
20..805						
/note="unnamed protein product"						
/codon_start=1						
/protein_id="CAC36466.1"						
/db_xref="GI:13619653"						
/translation="MAPPRGGAATAALDLTGHVILEASSVPLPERGNAVQRKG"						

CDS

AVDPDKRKEKAAAPRIITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMH  
IGGPEKNIIRRVYDAFNVLIALRVIAKEKEIRWGLSNTRYEKIKLEEVKELV  
NKIRNKALQLEIEKQFDLQNKIRNKQLEIYSAENVNGIRLFPVLVKTSRKARVEI  
ISDDSKFAHFEFNGAPFTLHDDLILLEGVRNSIGRAGRATLH"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,24e-97 Length: 1089  
Score: 1314.00 Matches: 260  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.62% Indels: 0  
DB: 6 Gaps: 0

US-10-088-830-2 (1-261) x AX100704 (1-1089)

QY 1 MetAlaProProArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGly 20  
Db ATGGCGCTCCCGCGCGGAGCTGCTGCGCGCGCTACCGCGCACTGGACCTGACCGG 79  
QY 21 ValHisIleLeuGluAlaSerValProProLeuProGluAlaGlyGlyAsnAlaVal 40  
Db GTGCACATTCGAAGCTTCCAGTGTCCCGCGCTTCCCGCAACGCGCGGTAATGCGGTC 139  
QY 41 GlnArgGlyGlyAlaValAspProAspLysAspArgLysGlyLysGlyLysAlaAlaPro 60  
Db CAAAGGAAGGGGCTGTGACCCGGATAAGATAGGAAGAGAGAGAGAGAGAGAGAGAG 199  
QY 61 ArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAla 80  
Db AGGATCACCCTGGGGCTCCCGGAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCC 259  
QY 81 LysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLysSerMet 100  
Db AAAGGAAGAACACATACATGAGTTGCAGACAAATTTATTTCAGAGCTGAAGTCCATG 319  
QY 101 AlaHisIleGlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPhe 120  
Db GCACATTTGGTCAAGGTTTGTAGAGAAATATTAGGCGGAGGTGTATGATGCTTTC 379  
QY 121 AsnValIleAlaLeuArgValIleAlaLysGluLysGlyLysGluIleArgTrpMetGly 140  
Db AACGTTCTCATTCGCTTCTGTTATTGCAAAAGAAAAAGAGAGATACGCTGATGGGC 439  
QY 141 LeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluValArgLysGluLeuVal 160  
Db CTTTCAATTTACAGATATGAAAAATAAGAAAGCTTGAGAGACTTCGTAAGAACTCGTC 499  
QY 161 AsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeu 180  
Db AACAGATTAGGAACAAAGAGGACCTCTCCAGAAATCGAAACAGTTTGAATGATCTC 559  
QY 181 GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGlyIle 200  
Db CAAAACATCAAGTTACGTACCAACAAACACATGAAAGCTCAGCAGAGAATGTTAATGCA 619  
QY 201 ArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleSer 220  
Db CGCTTTCATTCGTTATTTGGTCAAGACATCTAGAAAGCAAGGCTGAAATTCGATTTCA 679  
QY 221 AspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAsp 240  
Db GATGACTCGAAGTTTGCCTATTTGAGTTCAATGGTGACCATTCACATTTGCATGATGAT 739  
QY 241 LeuSerIleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260  
Db CTCTCAATCTCTGAGGGGTAAAGCGGTAAACAGCATAGGAAGAGCTGGCGCGCACCCCTT 759  
QY 261 His 261  
Db 800 CAC 802

RESULT 2

## TSP271917

LOCUS TSP271917 1083 bp mRNA linear PLN 16-DEC-2000  
DEFINITION Triticum sp. mRNA for DP protein (dp gene).  
ACCESSION AJ271917  
VERSION AJ271917.1 GI:11877790  
KEYWORDS DP gene; E2F dimerization partner.  
SOURCE Triticum sp.  
ORGANISM Triticum sp.

## REFERENCE

1 Ramirez-Parra,E. and Gutierrez,C.  
AUTHORS Characterization of wheat DP, a heterodimerization partner of the  
TITLE Plant E2F transcription factor which stimulates E2F-DNA binding  
JOURNAL FEBS Lett. 486 (1), 73-78 (2000)  
MEDLINE 20562804  
PUBMED 11108846

## REFERENCE

2 (bases 1 to 1083)  
AUTHORS Gutierrez,C.  
TITLE Direct Submission  
JOURNAL Submitted (03-FEB-2000) Gutierrez C., C.S.I.C.-U.A.M., Centro de  
Biologia Molecular, Cantoblanco, Madrid 28049, SPAIN

## FEATURES

source

1..1083 /organism="Triticum sp."

/mol\_type="mRNA"

/db\_xref="taxon:4569"

## gene

1..1083 /gene="dp"

<1..13

/genes="dp"

14..799

## CDS

/function="cell cycle E2F dimerization partner"

/codon\_start=1

/product="DP protein"

/protein\_id="CAC19034.1"

/db\_xref="GI:11877791"

/db\_xref="GOA:Q9FET1"

/translation="NAPPRGAAAAATAALDLTGVIHLEASSVPPLPERGGNAVORRG

IGGPEKNIIRRVYDAFNVLIALRVIAKEKEIRWGLSNTRYEKIKLEEVKELV

NKIRNKALQLEIEKQFDLQNKIRNKQLEIYSAENVNGIRLFPVLVKTSRKARVEI

ISDDSKFAHFEFNGAPFTLHDDLILLEGVRNSIGRAGRATLH"

801..1033

/gene="dp"

## 3' UTR

## ORIGIN

## Alignment Scores:

Pred. No.: 1,44e-96 Length: 1083  
Score: 1304.00 Matches: 258  
Percent Similarity: 99.23% Conservative: 1  
Best Local Similarity: 98.85% Mismatches: 2  
Query Match: 98.86% Indels: 0  
DB: 8 Gaps: 0

US-10-088-830-2 (1-261) x TSP271917 (1-1083)

QY 1 MetAlaProProArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGly 20

Db ATGGCGCTCCCGCGCGGAGCTGCTGCGCGCGCTACCGCGCACTGGACCTGACCGG 73

QY 21 ValHisIleLeuGluAlaSerValProProLeuProGluAlaGlyGlyAsnAlaVal 40

Db GTGCACATTCGAAGCTTCCAGTGTCCCGCGCTTCCCGCAACGCGCGGTAATGCGGTC 133

QY 41 GlnArgLysGlyAlaValAspProAspLysAspArgLysGlyLysGlyLysAlaAlaPro 60

Db CAAAGGAAGGGGCTGTTCACCGGATAAGATAGGAAGAGAGAGAGAGAGAGAGAG 193

QY 61 ArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAla 80





```
Db 194 AAAAAATCGGTTCTTTTACGACGGCGGGCGCAATCGGCAGAGAGGAGGGC 253
QY 44 GlyAlaValAspProAspLysAspArgLysLysGluLysAlaAlaProArgLysThr 63
Db 254 GGTGGG-----AACAGGAACGGGAAGAGAGAGACCAGGCGCGAGAGATCACC 304
QY 64 GlyTrpGlyLeuArgGluTrpSerLysLysLysValCysGluLysValGluAlaLysGlyArg 83
Db 305 GGGTGGGGCTCGGTGAGTTCAGACAGATAGTTTCTAAGAAAGTTCAGGCGCAAGGA 364
QY 84 ThrThrAsnGluValAlaAspGluLysLysLysLysLysLysLysLysLysLysLys 103
Db 365 ACCACATATAATAGGTTCGCGATGAGATTTTTCGCGAGCTGAAGTCCATTCAGCAGA 424
QY 104 GlyGlnGlyPheAspGluLysAsnLysLysLysLysLysLysLysLysLysLysLys 123
Db 425 GGTCTGGAGTTTGTATGAGAAGATATATTAGCGGAGGGTATATGATGCTTTCATGTGCTC 484
QY 124 IleAlaLeuArgValLysLysLysLysLysLysLysLysLysLysLysLysLysLys 143
Db 485 ATTGCAATCGTGTATTGCAAAAGATATAAAAGAGATAAAGTGGATGGGCTTTACTAAT 544
QY 144 TyrArgTrpGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 163
Db 545 TATAGATACGAAAGATACAGAAAGTTGGAGGAAGTTTCACAAAGAACTCATCACCAGATC 604
QY 164 ArgAsnLysLysAlaLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 183
Db 605 AAGAAATAGAGAAAGCTTCTCCAGGAATTTGAAGACAGTTTGTATGACCTTTCAGAAAT 664
QY 184 LysLeuArgAsnGlnThrLeuGluSerSerSerSerSerSerSerSerSerSerSer 203
Db 665 ACATTACGCAACCAGGCTAGTCAGAGCGCCAGCAGAAAGTTTAATGGCATCTCTTCG 724
QY 204 PheValLeuValLysThrSerArgLysAlaArgValGluLysLysLysLysLysLys 223
Db 725 TTCTTATTGATCAAGACATCCGAAAGCAAGGCGTGAATTTGAGATTTTCGGAAGATTCG 784
QY 224 LysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSer 243
Db 785 AGTTTGGACGGTTCAGCTTCAACGGTGCACCAATTCACCATGATGATGATCAATC 844
QY 244 LeuGluGlyValArgArgAsnSerLysGlyArgAlaGly 256
Db 845 CTTGAAGCCATCAGCGCTAACAAA---GGAAGAGCTGGC 880

RESULT 5
ATH294532 1158 bp mRNA linear PLN 15-DEC-2000
LOCUS Arabidopsis thaliana mRNA for DP-like protein (dpp gene).
DEFINITION Arabidopsis thaliana
ACCESSION AJ294532
VERSION AJ294532.1 GI:11125652
KEYWORDS DP-like protein; dpp gene.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Magyar, Z., Atanassova, A., De Veylder, L., Rombauts, S. and Inze, D.
TITLE Characterization of two distinct DP-related genes from Arabidopsis
thaliana
JOURNAL FEBS Lett. 486 (1), 79-87 (2000)
MEDLINE 20562805
PUBMED 11108847
REFERENCE 2 (bases 1 to 1158)
AUTHORS Magyar, Z.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2000) Magyar Z., Department Plantgenetica,
Laboratorium voor Genetica, Universiteit Gent, K.L. Ledeganckstraat
35., B-9000, Gent, BELGIUM
FEATURES
Location/Qualifiers
1..1158 source
```

```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="Columbia"
/db_xref="taxon:3702"
1..1158
/genes="dpp"
/genes="dpp"
/function="cell cycle regulation"
/codon_start=1
/product="DP-like protein"
/protein_id="CAC15484.1"
/db_xref="GI:11125653"
/db_xref="GOA:O9FNV2"
/translation="MTTGSNSHHHNSNNNNPSTRSMGTAVSQSVSTSGSMGS
PSRSQTLIVVTSTSDTTFQRLNLDIOGDDAGSQAGSGVKKKKQGRAGRPDKGR
GQRQSMKCKEYSGRTTYNEVDELVAEFAFPNNDGTSPQOQYQRRARRRVYD
ALNVLAMMDIIISKDKKEIQWRGLPTSLDIEELKNERLSLRNRIEKKTKYSGLEEQ
YVGLQNLIORNEHLSSYSGVALPFLVOTRPHATVEVISEDMLVHDFNST
PFELHDDNFVLTKMFCDDPQQPQNGRNSQLVCHNPTPENPKNGSTGPTQLDMYE
THLQSQHQHSQLQIIPMPETNNVTSSADTAPVKSPSLPGINSSMKPEN"

ORIGIN
Alignment Scores: 8.64e-31 Length: 1158
Pred. No.: 492.00 Matches: 104
Score: 66.06% Conservative: 40
Percent Similarity: 47.71% Mismatches: 68
Best Local Similarity: 37.30% Indels: 6
Query Match: 8 Gaps: 2
DB: 2
US-10-088-830-2 (1-261) x ATH294532 (1-1158)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
Db 220 GGTGATGATGCTGCTTCTCAAGGAGCTTCTGTTTAAAGAAAGAAAGAGGAGGACGCT 279
QY 57 AlaAlaAlaProArgLysLysGlyTrpGlyLeuArgGluTrpSerLysLysLysLys 76
Db 280 GCGGCTGGTCCAGATTAAGACTGGAAGAGACTACGTCATTTAGTATGAAAGTTTGTGAA 339
QY 77 LysValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThr 96
Db 340 AAGGTGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 399
QY 97 LeuLysSerMetAlaHisLysLysLysLysLysLysLysLysLysLysLysLysLys 111
Db 400 TTTGCACATTCCAAATAACGATGGAACATCCCTCGATCAGCAACAGATGATGAGAAAAC 459
QY 112 IleArgArgArgValTrpAspAlaPheAsnValLysLysLysLysLysLysLysLys 131
Db 460 ATAAGCAGAGAGATATATGATGCTTTAAACGCTCTCATGGCTATGATATAATATCAAG 519
QY 132 GluLysLysGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 151
Db 520 GATAAAAGAAAGAAATTCATGGAGAGCTCTCTCCGACAGAGCTTAAAGCGACATTAAGAA 579
QY 152 LeuGluGluValArgLysGluLysValAsnLysLysLysLysLysLysLysLysLys 171
Db 580 TTAAGAACGACACGACTCTCACTTAGACAGAAATTCAGAAAGAAAGAAAGAAAGAA 639
QY 172 GluLysLysGlnPheAspAspLeuGlnAsnLysLysLysLysLysLysLysLysLys 191
Db 640 GAACCTGGAAGAAACAATATGTAGCGCTTCAGAAATCTCATCAGAGAAATGAGCACTATAT 699
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
Db 700 AGCTCAGGAATGCTCCAGTGGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 759
QY 211 ArgLysAlaArgValGluLysLysLysLysLysLysLysLysLysLysLysLysLys 230
Db 760 CCTCAGCAACAGTAGAAGTGGAGATATCAGAAAGATATCAGAGATATGAGCTCGTCACTT 819
```



QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248  
 Db 820 AACAGCACTCCATTGAGCTCCACGAGGACAATTTGCTCCTCAAGACTATGAAG 873

RESULT 6  
 AY086018 1618 bp mRNA linear PLN 14-APR-2003  
 LOCUS Arabidopsis thaliana clone 206791 mRNA, complete sequence.  
 DEFINITION Arabidopsis thaliana (thale cress)  
 ACCESSION AY086018.1 GI:21404728  
 VERSION FLI CDNA.  
 KEYWORDS Arabidopsis thaliana  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi;  
 1 (bases 1 to 1618)  
 Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,  
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.  
 Full-length messenger RNA sequences greatly improve genome  
 annotation  
 Genome Biol. 3 (6), RESEARCH0029 (2002)  
 22088475  
 MEDLINE 12093376  
 PUBMED  
 REFERENCE 2 (bases 1 to 1618)  
 BROVER,V., TROUKHAN,M., ALEXANDROV,N., LU,Y.-P., FLAVELL,R. and  
 FELDMANN,K.  
 Full-length cDNA from Arabidopsis thaliana  
 Unpublished  
 3 (bases 1 to 1618)  
 BROVER,V., TROUKHAN,M., ALEXANDROV,N., LU,Y.-P., FLAVELL,R. and  
 FELDMANN,K.  
 Direct Submission  
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
 Malibu, CA 90265, USA  
 This clone sequence is one of 5,000 Ceres full-length cDNAs made  
 available to TIGR and Genbank. The following quality assessment of  
 this set was done by comparison with known proteins: two percent of  
 the clones are estimated to be 5'-truncated; less than one percent  
 are 3'-truncated; approximately two percent represent alternative  
 splice variants, including unspliced introns and spliced exons; one  
 percent may contain premature stop codons; five percent may have  
 frame shifts in a coding region. A sequence is considered to be  
 5'-truncated if it lacks the translation initiation start (ATG). A  
 sequence is considered to be 3'-truncated if it lacks the  
 C-terminal end of the encoded protein. Please note that these cDNA  
 sequences are derived from the Ws or Laer ecotypes and therefore  
 may contain polymorphisms when compared to sequences from Col-0.  
 GenSet carried out the library production and sequencing of the  
 full-length clones. Ceres, Inc. carried out the clustering of the  
 5' sequences, selection of clones, and sequence assembly.

FEATURES  
 source Location/Qualifiers  
 1..1618  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clones="206791"  
 249..1406  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA063227.1"  
 /db\_xref="GI:21534147"  
 /translation="MTTSGNSNNHHNNNNNNNNPSTRGWGTAVSQSVSTSSMGS  
 PSSRSSEITVTTSASDTTIFQLNNLDIQDAGSQGAGVKKKKRQRAAGPKTKR  
 GLRQFSKVCESKGRITTYNEVADELVAEFLPNDGTSPOQOQYDKNIIRRLVYD  
 ALNYLMAIDTISKDKETOMRGLPRTSLSDIEELKNERLSLRNRIEKKTYASLEQ  
 YVGLQNLQIQRNHLSSGNAPSGGVALPFLIVQTRPHATVEVEISEDMOIAVDFDNEST  
 PFLHDNFVLTWKRFCDQPPQPNQGNNSQLVCHNFTPENPKGPTSGTGPQLDMYE  
 THLSQOQHQQHSLQLIIPMPETNNVTSSAUTAPVKPSFLPRIMMSMKPEN"

CDS  
 249..1406  
 /product="unknown"  
 /protein\_id="AA063227.1"  
 /db\_xref="GI:21534147"  
 /translation="MTTSGNSNNHHNNNNNNPSTRGWGTAVSQSVSTSSMGS  
 PSSRSSEITVTTSASDTTIFQLNNLDIQDAGSQGAGVKKKKRQRAAGPKTKR  
 GLRQFSKVCESKGRITTYNEVADELVAEFLPNDGTSPOQOQYDKNIIRRLVYD  
 ALNYLMAIDTISKDKETOMRGLPRTSLSDIEELKNERLSLRNRIEKKTYASLEQ  
 YVGLQNLQIQRNHLSSGNAPSGGVALPFLIVQTRPHATVEVEISEDMOIAVDFDNEST  
 PFLHDNFVLTWKRFCDQPPQPNQGNNSQLVCHNFTPENPKGPTSGTGPQLDMYE  
 THLSQOQHQQHSLQLIIPMPETNNVTSSAUTAPVKPSFLPRIMMSMKPEN"

ORIGIN  
 Alignment Scores:

Pred. No.: 1..31e-30 Length: 1618  
 Score: 492.00 Matches: 104  
 Percent Similarity: 66.06% Conservative: 40  
 Best Local Similarity: 47.71% Mismatches: 68  
 Query Match: 37.30% Indels: 6  
 Gaps: 2

US-10-088-830-2 (1-261) x AY086018 (1-1618)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56  
 Db 468 GGTGATGATGCTGGTCTCTCAAGAGAGCTTCTGGTGTAAAGAAGAAGAGGGGAGCGGT 527

QY 57 AlaAlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGlu 76  
 Db 528 GCGGCTGCTCCAGATAACACTGGAAGAGAGTACCTCAATTAGTATGAAGATTGTGTA 587

QY 77 LysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGlu 96  
 Db 588 AAGTGGAAAGCAAGGAAGGACACTTACATGAGGTTCGACAGAGCTTGTCTGTA 647

QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111  
 Db 648 TTTGCACCTTCCAAATAACGATGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAAC 707

QY 112 IleArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131  
 Db 708 ATAAGCAGCAAGAGTATATGATGCTTTAAACGTCCTGAGCTATGATATATATCCAG 767

QY 132 GluLysLysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151  
 Db 768 GATAAAAGAAATTCATGGAGAGGTCTTCTCGCAAGCTTAAGCCGACATTCGAAGA 827

QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlu 171  
 Db 828 TTAAGAAGCAAGCAGCTCTCCTTAGGACAGAAATTGAGAAAGAACTGCATATCCCAA 887

QY 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191  
 Db 888 GAATCGAAGAACATATGATGAGCTTCAATCTGATACAGAGAAATGACGACTTATAT 947

QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuPropheValLeuValLysThrSer 210  
 Db 948 AGCTCAGGAAATGCTCCAGTGGCGGTGTCTCTCTCTTTTCTTTTCTTCCAGACTCGT 1007

QY 211 ArgLysAlaArgValGluIleSerAspAspSerLysPheAlaHisPheGluPhe 230  
 Db 1008 CTTACGCAACAGTAGAAGTGGAGATATCAGAAGATATGAGCTCGTCTGATTTGATTTC 1067

QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248  
 Db 1068 AACAGCACTCCATTGAGCTCCACGAGGACAATTTGCTCCTCAAGACTATGAAG 1121

RESULT 7  
 AF181998 1640 bp mRNA linear PLN 31-DEC-2003  
 LOCUS Populus tremula x Populus tremuloides transcription factor DPI  
 DEFINITION (DPI) mRNA, complete cds.  
 ACCESSION AF181998  
 VERSION AF181998.1 GI:33338105  
 KEYWORDS  
 SOURCE  
 ORGANISM Populus tremula x Populus tremuloides  
 Populus tremula x Populus tremuloides  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
 1 (bases 1 to 1640)  
 Bhalerao,R.P. and Sandberg,G.  
 A DP-1 homolog clone from hybrid Aspen  
 Unpublished  
 2 (bases 1 to 1640)  
 Bhalerao,R.P. and Sandberg,G.  
 Direct Submission



Qy	188	GlnThrIeuGluSerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValIeu	208
	:::		:::
Db	843	GAACAACTGTACAGCTCAGGAAATGCTCCTAGTGGTGTGTGCTGCTTTTATTCTG	902
Qy	207	ValIysThrSerArgIysAlaArgValGluIleGluIleSerAspSerLysPheAla	226
	:::		:::
Db	903	GTCCAGACACGCCCTCATGCAACTGTTGAAGTGGAGATATCAGAAGATATGCAGCTGGTT	962
Qy	227	HisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGly	246
	:::		:::
Db	963	CACTTTGATTTTATAGCACTCCCTTCGAGCTCATGACGATATTAACGTTCTCAAGCA	1022
Qy	247	ValArg	248
	:::		:::
Db	1023	ATGAAA	1028
RESULT	8		
LOCUS	AY307373		
DEFINITION	Populus tremula x Populus tremulooides DP1 (DP1) mRNA	linear	PLN 24-JUN-2003
ACCESSION	AY307373		
VERSION	AY307373.1	GI:32187096	
KEYWORDS			
SOURCE	Populus tremula x Populus tremulooides		
ORGANISM	Populus tremula x Populus tremulooides		
REFERENCE	Populus tremula x Populus tremulooides		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus. 1 (bases 1 to 1640)		
TITLE	Espinosa-Ruiz,A., Saxena,S., Schmidt,J., Mellerowicz,E., Bako,L.S. and Bhallerao,R.P.		
JOURNAL	Differential stage specific regulation of cyclin dependent kinases during cambial dormancy in hybrid aspen		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1640)		
TITLE	Espinosa-Ruiz,A. and Bhallerao,R.P.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (27-MAY-2003) Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences (SLU), Umea		
JOURNAL	907 34, Sweden		
FEATURES	Location/Qualifiers		
source	1..1640		
	/organism="Populus tremula x Populus tremulooides"		
	/mol_type="mRNA"		
	/db_xref="taxon:47664"		
gene	1..1640		
	/genes="Dp1"		
CDS	156..1217		
	/genes="Dp1"		
	/codon_start=1		
	/product="Dp1"		
	/protein_id="AAP73785.1"		
	/db_xref="GI:32187097"		
	/translation="MVAGGAGHLEDDGRHPSSRRGGGGATTGSWVGQSVSTGSGVGSPPSRSEHAMATPADSDTFLNLHLDIHADDAATDAANKKKKGQAVAGDKSGRGLRPSIKVCEKVESKGTITYNEVADLVAEPADPSNSYSTPQQOYDEKNIRRVRYDALNLMALDIIKDKKEIKWGLPTSLSDIEELKVERLGLRNFKKKAYLQLELEQFVGLQNLQRNQLYSSGNAPSGVSLPILQTRPHATVEVEISDMQLVHDFNSTFPELHDDNLYLKAMKFCRPSQDGMAPNPPADGEGSSMSMYQPQILLASPSPTNTPVRHPTSPPLPGIIRKARVKNH"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,39e-29	Length:	1640
Score:	476.50	Matches:	112
Percent Similarity:	59.16%	Conservative:	43
Best Local Similarity:	42.75%	Mismatches:	76
Query Match:	36.13%	Indels:	31
DB:	8	Gaps:	6
US-10-088-830-2 (1-261) x AY307373 (1-1640)			



Db 498 GGATATTATCTAAGAGTAAAGGAAATTCAGTGAAGGCTTGCTCGG----- 549  
 QY 146 rGluLysIleLysLysLeuGluValArgLysGluLeuVal-----AsnLysIle 163  
 Db 550 -ACAAGTATGAGGATGTTGAAGAAATTGAAGACAGAGTATCGGACTGAAGGTAGGAT 608  
 QY 163 eArgAsnLysLysAlaLeuLeuGlnGluLysGlnPheAspAspLeuGlnAsnIle 183  
 Db 609 CGACAAGAAAATGCATATTGTCAGGAGTTAGAAGATCAATTGCTAGGCTTCAAACTT 668  
 QY 183 eLysLeuArgAsnGlnThrLeuGlnSerSerAlaGluAsnValAsn---GlyLeuArgLe 202  
 Db 669 GCACACCGCAAGCAGAGCTTATGCTTTCAGGAAATGCTCTTCAGGAGGAGTGGCAT 728  
 QY 202 uProPheValLeuValLysThrSerArgLysAlaArgValGluLeuLeuLysSerAsp 222  
 Db 729 GCCATTATATTGGTCGACACAGCTCTCATGCTACAGTAGAAGTGGAGATATCAGAAGA 788  
 QY 222 pSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSe 242  
 Db 789 TATGCAGCTGGTCATTTTCATTTTCAATAGCACTCCATTTGAACATGCATGACGATTCCTT 848  
 QY 242 rIleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArg 257  
 Db 849 TGTAAGTAA-----GCATTGGGGTTCCTCGCAAA 879

RESULT 10  
 LOCUS AX449314 1131 bp DNA linear PAT 03-JUL-2002  
 DEFINITION Sequence 33 from Patent WO0185946.  
 ACCESSION AX449314  
 VERSION AX449314.1 GI:21698061

KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1  
 AUTHORS Inze,D., Boudolf,V., de Veylder,L., Acosta,J.A. and Magyar,Z.  
 TITLE Nucleic acid molecules encoding plant cell cycle proteins and uses  
 JOURNAL Patent: WO 0185946-A 33 15-NOV-2001;  
 CropDesign N.V. (BE)

FEATURES  
 source Location/Qualifiers  
 1..1131  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.74e-29 Length: 1131  
 Score: 466.50 Matches: 101  
 Percent Similarity: 63.76% Conservative: 38  
 Best Local Similarity: 46.33% Mismatches: 64  
 Query Match: 35.37% Indels: 15  
 DB: 6 Gaps: 3

US-10-088-830-2 (1-261) x AX449314 (1-1131)

QY 37 GlyAsnAlaValGlnArgLysGlyValAlaValAspProAspLysAspArgLysLysGluLys 56  
 Db 220 GGTGATGATCGGTCTCAAGAGCTTCTGGTTTAAGAAGAGAGGAGGAGCGCT 279  
 QY 57 AlaAlaAlaProArgIleThrGlyLysLeuArgLysSerLysIleValCysGlu 76  
 Db 280 CGCGCTGCTCCAGATAAGACTGCAAGAGGACTAGCTCAATTAGTATGAAGTTCTGAA 339  
 QY 77 LysValGluAlaLysGlyArgThrThrThrAsnGluValAlaAspGluLeuLysSerGlu 96  
 Db 340 AAGGTGGAAGCAAGGAGGACCACTTACATGAGGTTGCAGACGAGCTTGTGTCTGAA 399

QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111  
 Db 400 TTTGCACTTCCAAATACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 459  
 QY 112 IleArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131  
 Db 460 ATAAGACGAAGAGTATGATGCTTTAAACGCTCCTCATGCTATGATATATATCAAG 519  
 QY 132 GluLysLysGluLysLeuArgTyrMetGlyLysSerAsnTyrArgTyrGluLysIleLys 151  
 Db 520 GATAAAAAAGAAATTCATGAGAGGCTTCTTCGGAACAGCTTAAGCGCATGAAGAA 579  
 QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGln 171  
 Db 580 TTAAAGACGAAGCACTCTCATTAGGACAGAAATTCAGAGAAACTGCATATCCCAA 639  
 QY 172 GluLysLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191  
 Db 640 GAACTGGAAGACAA-----AGAAATGAGCACTTATAT 672  
 QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210  
 Db 673 AGCTCAGAAATGCTCCAGTCGCGGTGTCTTCCTTTATCTCTGTCAGACTCGT 732  
 QY 211 ArgLysAlaArgValGluLysLeuHisAspAspSerLysPheAlaHisPheGluPhe 230  
 Db 733 CTTACGCAACAGTAGAGTAGAGATATCAGAAGATATGCGCTCGTGCATTTGATTC 792  
 QY 231 AsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248  
 Db 793 AACAGCACTCATTGAGCTCCAGCAGCAATTTTGTCTCTCAAGACTATGAAG 846

RESULT 11  
 LOCUS AY224589 1140 bp mRNA linear PLN 05-FEB-2004  
 DEFINITION Oryza sativa (japonica cultivar-group) isolate 31182 E2F  
 dimerization factor mRNA, complete cds.  
 ACCESSION AY224589  
 VERSION AY224589.1 GI:29371982

KEYWORDS Oryza sativa (japonica cultivar-group)  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 1140)  
 AUTHORS Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,  
 Ellero,C., Goff,S.A. and Glazebrook,J.  
 TITLE Identification of rice (Oryza sativa) proteins linked to the  
 cyclin-mediated regulation of the cell cycle

JOURNAL Plant Mol. Biol. 53 (3), 273-279 (2003)  
 MEDLINE 23111120  
 PUBMED 14750518  
 REFERENCE 2 (bases 1 to 1140)  
 AUTHORS Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,  
 Ellero,C., Goff,S. and Glazebrook,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JAN-2003) Torrey Mesa Research Institute, Syngenta  
 Research and Technology, 3115 Merryfield Row, San Diego, CA 92121,  
 USA

FEATURES  
 source Location/Qualifiers  
 1..1140  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /isolates="31182"  
 /db\_xref="taxon:39947"  
 /note="derived from yeast two-hybrid experiments"  
 1..1140  
 /codon\_start=1  
 /product="E2F dimerization factor"  
 /protein\_id="AA072709.1"  
 /db\_xref="GI:29371983"

CDS





```

Db      619 GCAGTAGTCGAATCGAGATTCTGAGATATGCAACTGTACACCTCGACTTCATATAGC 678
QY      233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245
Db      679 ACACCTTTCTCGTCCATGATGCTTACATTTTGAAA 717

RESULT 14
LOCUS   ATH319027
DEFINITION Arabidopsis thaliana partial mRNA for E2F dimerisation partner
ACCESSION AJ319027
VERSION   AJ319027.1 GI:18447785
KEYWORDS  dp2a gene; E2F dimerisation partner protein.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS   Rossignol,P.
TITLE      E2F family transcription factors: AtE2F-a and AtDP-a, induce
JOURNAL    Arabidopsis leaf cells to re-enter S phase
REFERENCE
AUTHORS   Bergounioux,C.
TITLE      Unpublished
JOURNAL    2 (bases 1 to 960)
REFERENCE
AUTHORS   Direct Submission
TITLE      Submitted (29-AUG-2001) Bergounioux C., UMR 8618, Cnrs, IBP bat 630
JOURNAL    Universite Paris-Sud, 91405 Orsay, FRANCE
FEATURES
source
1..960
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Colombia"
/db_xref="taxon:3702"
/tissue_type="buds"
/dev_stage="flowering"
1..960
/gene="dp2a"
<1..>960
/gene="dp2a"
79..957
/gene="dp2a"
/functions="transcription"
/codon_start=1
/product="E2F dimerisation partner protein (DP2a)"
/protein_id="CAC87459.1"
/db_xref="GI:18447786"
/db_xref="GOA:Q9FNV3"
/db_xref="UniProt/TrEMBL:Q9FNV3"
/translators="MSMEMELFVTPKQROHVSVEKTPVRRLIVDDDDSEIGSEKK
GQRTSGGGLRQFSVMVCQKLEAKITTYKEVADEIISDPATIKQNAEKLNEVNE
KNTRRVYDALNFMALDIARQKEITRWKGLPTCKDVEEVMDBRNKVMSSVORKA
AFUKELEKVSLESLMSRNQVMVKTQGPAGFTLPFILLETNPRAVVEIETSDMQ
LVHLDNSTPFSVHDDAYILKLMQEQKQEQNRVSSSSSTHHQSHSSSSSSCIAS
GTSGPVCWNSGSDTR"

ORIGIN
Alignment Scores:
Pred. No.: 9,48e-26 Length: 960
Score: 428.50 Matches: 88
Percent Similarity: 64.32% Conservative: 49
Best Local Similarity: 41.31% Mismatches: 69
Query Match: 32.49% Indels: 7
DB: 8 Gaps: 2

US-10-088-830-2 (1-261) x ATH319027 (1-960)

QY      40 ValGlnArgLysGlyAlaValAspProAspLysAspArgLysGlyLysAlaAla 59
Db      157 GTGAGAGGAATGATTGTTGTTGATGATCTCTGAAATGGATCAGAGAGAAAGGCCAA 216
QY      60 ProArgIleThrGlyTTPGlyLeuArgGluTyrSerLysIleValCysGluLysValGlu 79

```

```

Db      217 TCAGACTTCTGGAGCGGGCTTCGTCAATCACTGTTATGTTTGTTCAGAA 276
QY      80 AlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99
Db      277 GCCAAGAAGATAACTACTTACAAAGGAGGTTGCAGACGAAATATTATTTCAGATTTGCACA 336
QY      100 MetAlaHisIleGlyGln-----GlyPheAspGluLysAsnIleArg 113
Db      337 ATTAAAGCAAAACGACAGAGAGCGCTTTGAATGAAAATGAGTACAAATGAGAAGAACATAAGG 396
QY      114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db      397 CGAGAGCTCAGATCGCTCAATGTTTCATGGCGTTGGATATTATTTCAGAGGGATAAA 456
QY      134 LysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGlyLysIleLysLysLeuGlu 153
Db      457 AAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAGAGATGTGGAAGAGTCAAG 516
QY      154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGluIle 173
Db      517 ATGGATCGTAATAAAGTTAAGACGAGTGTGCAAAAGAGCGTCTTTCTTAAAGAGTTG 576
QY      174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db      577 AGAGAAAGGCTCTCAAGTCTTCAGAGCTTATGTCGAGAAATCAAGAGATGGTTGTGAAG 636
QY      194 AlaGluAsn---ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212
Db      637 ACTCAAGGCCACGACGAGAGGATTTACCTTACCATTCATCTACTTACGACAAACCCCTCAC 696
QY      213 AlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPheAsnGly 232
Db      697 GCAGTAGTCGAATCGAGATTTCTGAGATATGCAACTGTGTACACCTCGACTTCATATAGC 756
QY      233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245
Db      757 ACACCTTTCTCGTCCATGATGATGCTTACATTTTGAAA 795

RESULT 15
LOCUS   AX449297
DEFINITION Sequence 16 from Patent WO0185946.
ACCESSION AX449297
VERSION   AX449297.1 GI:21698044
KEYWORDS  Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS   Inze,D., Boudolf,V., de Veylder,L., Acosta,J.A. and Magyar,Z.
TITLE      Nucleic acid molecules encoding plant cell cycle proteins and uses
JOURNAL    Patent: WO 0185946-A 16 15-NOV-2001;
therefor
CropDesign N.V. (BE)
FEATURES
source
1..1114
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 1.14e-25 Length: 1114
Score: 428.50 Matches: 88
Percent Similarity: 64.32% Conservative: 49
Best Local Similarity: 41.31% Mismatches: 69
Query Match: 32.49% Indels: 7
DB: 6 Gaps: 2

US-10-088-830-2 (1-261) x AX449297 (1-1114)

```

```
QY 40 ValGlnArgLysGlyAlaValAspProAspLysArgLysLysGluLysAlaAla 59
Db |||:||||| ||||| ||| :|||
92 GTGAGAAGGAATTGATTGATGATTCGAAATTCGATCAGAGAAGAGGCCAA 151
QY 60 ProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGlu 79
Db |||:||||| ||||| ||| :|||
152 TCAAGAACCTTCGAGGCGGCTTCGTCATTCAGTGTATGTTATGTTCTCAGAAGTTGAA 211
QY 80 AlaLysGlyArgThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GCCAGAAGATCACTACTACAGGAGGTTGCAGACGAATATATTCAGATTTGCCACA 271
QY 100 MetAlaHisIleGlyGln-----GlyPheAspGluLysAsnIleArg 113
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
272 ATTAACCAAAACGCACAGAGCCCTTTGAATGAAATCAGTACATCAGAGAACAATAAGG 331
QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 CGGAGAGTCTACGATCGCTCAATGTGTTCATGCGCTGCGATATTATTCAGGGATAAA 391
QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 AGGAAATCCGGTGAAGAGACTTCCTATTACCTGCAAAAGGATGTGGAAGAGTCAAG 451
QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle 173
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
452 ATGGATCGTAAATAAAGTTATGACGCTGTGCAAAAGAGGCTGCTTTCTTAAAGAGTTG 511
QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
512 AGAGAAAAGGCTCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGTTGTGAAG 571
QY 194 AlaGluAsn---ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
572 ACTCAAGGCCCCAGCAGAGGATTTACCTTACCATTATTCTACTCTGAGACAAACCTCAC 631
QY 213 AlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPheAsnGly 232
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 GCAGTAGTCGAATCGAGATTCTGAGATATGCAACTTGTACACCTCGACCTCAATAGC 691
QY 233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
692 ACACCTTTCTCGTCCATGATGCTTACATTTTGAAA 730
```

Search completed: February 6, 2005, 23:22:05  
Job time : 4661 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 6, 2005, 18:50:47 ; Search time 583 Seconds  
(without alignments)

2650.176 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlh  
-Q/cgn2\_1/USPTO.spool/US10088830/runat\_04022005\_122602\_29549/app\_query.fasta\_1.455  
-DB=N Geneseq 16Dec04 -QFWT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us10088830 @CGN 1.1 470 @runat\_04022005\_122602\_29549 -NCPU=6 -ICPU=3  
-NO\_WRAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:\*  
1: Geneseqn19808:\*  
2: Geneseqn19908:\*  
3: Geneseqn20008:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001be:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1314	99.6	1089	4 AAF80144	Nucleotide
2	890.5	67.5	885	12 ADQ36842	Adq36842 Cell prol
3	492	37.3	1158	12 ADO63082	Ado63082 Transcrip
4	492	37.3	1158	12 ADO62142	Ado62142 Transcrip
5	492	37.3	1618	3 AAC39795	Aac39795 Arabidops

6	482.5	36.6	1443	12	ADO62852	Ado62852 Transcrip
7	476.5	36.1	1153	12	ADO63626	Ado63626 Transcrip
8	471.5	35.7	1041	12	ADQ36844	Adq36844 Cell prol
9	466.5	35.4	1131	6	AAS96304	Aas96304 Arabidops
10	466.5	35.4	1140	12	ADQ36854	Adq36854 Cell prol
11	463	35.1	1245	12	ADO63363	Ado63363 Transcrip
12	461	35.0	900	3	AAA59702	Aaa59702 DNA encod
13	456	34.6	1442	6	AAS96332	Aas96332 Arabidops
14	455	34.5	1193	2	AZ34580	Aaz34580 Corn DP-2
15	428.5	32.5	1114	6	AAS96287	Aas96287 Arabidops
16	428.5	32.5	1189	6	AAS96319	Aas96319 Arabidops
17	428.5	32.5	1267	12	ADO63084	Ado63084 Transcrip
18	428.5	32.5	1267	12	ADO62144	Ado62144 Transcrip
19	428.5	32.5	1274	10	ADH59559	Adh59559 Dimerisat
20	428.5	32.5	1274	10	ACC45100	Acc45100 Arabidops
21	399.5	30.3	2320	11	ADI32130	Adi32130 Human cDN
22	399.5	30.3	4359	12	ADN05729	Adn05729 Antipeori
23	396.5	30.1	1396	5	AAS92244	Aas92244 DNA encod
24	394	29.9	1266	6	ABL65866	Ab165866 Lung canc
25	394	29.9	2783	13	ACN38014	Acn38014 Tumour-as
26	394	29.9	2968	5	ABV26838	Abv26838 Human pro
27	394	29.9	2968	5	ABV20992	Abv20992 Human pro
28	389	29.5	1385	2	AAX76730	Aax76730 Murine DP
29	387	29.3	2968	11	ACN88986	Acn88986 Breast ca
30	378.5	28.7	1700	2	AAT10582	Aat10582 Transcrip
31	376	28.5	870	6	AAS96410	Aas96410 Arabidops
32	375.5	28.5	2203	4	ABL08071	Ab108071 Drosophil
33	374.5	28.4	692	10	AD59591	Ad59591 Human tra
34	374.5	28.4	1440	10	ADK67024	Adk67024 Gene #114
35	374.5	28.4	1440	11	ADI32073	Adi32073 Human cDN
36	374.5	28.4	1440	12	ADM78581	Adm78581 Human tra
37	374.5	28.4	1440	12	ADP13337	Adp13337 Renal cel
38	374.5	28.4	1440	13	ADR25496	Adr25496 Breast ca
39	374.5	28.4	2394	12	ADO19868	Ado19868 Human PRO
40	374.5	28.4	2607	4	AAH33503	Aah33503 Human col
41	372.5	28.2	1700	2	AAQ65344	Aaq65344 Transcrip
42	370	28.1	642	13	ADR65325	Adr65325 Cotton CD
43	361.5	27.4	962	12	ADO63216	Ado63216 Transcrip
44	361.5	27.4	2673	6	AAS94979	Aas94979 Human DNA
45	361	27.4	1317	12	ADO62851	Ado62851 Transcrip

#### ALIGNMENTS

RESULT 1  
AAF80144  
ID AAF80144 standard; DNA; 1089 BP.  
XX  
AC AAF80144;  
XX  
DT 11-JUN-2001 (first entry)  
XX  
DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.  
XX  
KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;  
KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;  
KW ss.  
XX  
OS Triticum monococcum.  
XX  
FH Key Location/Qualifiers  
CDS 20..805  
FT /\*tag= a  
FT /product= "E2F-dimerisation partner (DP) protein"  
XX  
PN WO200121644-A2.  
XX  
PD 29-MAR-2001.  
XX  
PF 25-SEP-2000; 2000WO-EF009325.  
XX  
PR 24-SEP-1999; 99ES-00002127.  
PR 11-NOV-1999; 99ES-00002474.



Query Match: 67.51% Indels: 41  
DB: 12 Gaps: 4

US-10-088-830-2 (1-261) x ADQ36842 (1-885)

QY 1 MetAlaProArgGlyGlyAlaAlaAlaAlaThraAlaLeuAspLeuThrGly 20  
Db 1 ATGGCGCCTCCCTCGCGCGATCGCGCGCTCCCTCCCGCGCGCGCTGGCCAAAC 60  
QY 21 ValHisIleLeuGluAlaSerValProLeuProGlu 34  
Db 61 CTGCTCATTCGGAAGCGCGCTCCCTCCCGTCCCGAACGGTACCGCCATTCGG 120  
QY 34 34  
Db 121 CCCTGTACTTCTGATTCCTTGTCCCAATCTCTAGGGAAGGACGATATTCCTCCCA 180  
QY 35 35  
Db 181 AAAAAATCGTTCTTTACGCGCGCGCGCGCAATCGCGACGAGGAGGAGGC 240  
QY 44 GlyAlaValAspProAspLysAspArgLysGlyLysAlaAlaAlaProArgIleThr 63  
Db 241 GGTGCG-----AACAGAACCGGAAGAGAGAACCGCGCGCGACGAGGATCAC 291  
QY 64 GlyTyrGlyLeuArgGluTyrSerLysIleValCysGlyLysValGluAlaLysGlyArg 83  
Db 292 GGTGGGGCTCGTGGTTCACAGATAGTTCTTAAGAAAGTTAGGCGCAAGGAAGA 351  
QY 84 ThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHisIle 103  
Db 352 ACCACATATAATGAGGTTCGATGAGATTTTTCGCGAGCTGAAGTCCATACGAGAAC 411  
QY 104 GlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeu 123  
Db 412 GGTCTGAGTTGATGAGAGATATTTAGCGGAGGTATATGATGCTTTCAATGTGCTC 471  
QY 124 IleAlaLeuArgValIleAlaLysGluLysLysGluIleArgTyrMetGlyLeuSerAsn 143  
Db 472 ATTGCAATTCGTATTTCGAAAGATTAAGAGATTAAGTGGATGGCCCTTACTAAT 531  
QY 144 TyrArgTyrGluLysIleLysLysLeuGluValArgLysGluLeuValAsnLysIle 163  
Db 532 TATAGATACGAAAGATACAGAGTTGGAGGAAGTTTCAAAAGAACTCATCAGGATC 591  
QY 164 ArgAsnLysLysAlaLeuGluGlnGluIleGluLysGlnPheAspLeuGlnAsnIle 183  
Db 592 AAGAATAAGAAAGAGCTTCTCCAGGAATTTGAAAGCAGTTTGTATGACCTTCAGAAATAT 651  
QY 184 LysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuPro 203  
Db 652 ACATTCAGCAACGAGCTAGAGCGCGAGAGAGTGTATGATGCTTCCTTCCTCG 711  
QY 204 PheValLeuValLysThrSerArgLysAlaArgValGluIleGluLysSerAspSer 223  
Db 712 TTCTTATTCATCAAGACATCCCGAAAGCAAGGTGAAATTCAGATTCGGAAGATTCG 771  
QY 224 LysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIle 243  
Db 772 AAGTTTTCAGGTTTCGACTTCAACGGTGCACCATTCACCATGATGATGATCAATC 831  
QY 244 LeuGluGlyValArgAsnSerIleGlyArgAlaGly 256  
Db 832 CTTGAAGCCATCAGCGGTAAACAAA---GGAAGAGCTGGC 867

RESULT 3  
ID ADQ63082  
XX ADQ63082 standard; DNA; 1158 BP.  
AC ADQ63082;  
XX  
DT 15-JUL-2004 (first entry)  
XX

Transcription factor G2981 coding sequence, SEQ ID 1549.  
Plant; transcription factor; transgenic plant; abiotic stress tolerance; osmotic stress tolerance; cold tolerance; heat tolerance; low nitrogen tolerance; low phosphate tolerance; fungal disease; glyphosate resistance; flowering; fertility; seed development; ds. Arabidopsis thaliana.  
WO2004031349-A2.  
15-APR-2004.  
18-SEP-2003; 2003WO-US030292.  
18-SEP-2002; 2002US-0411837P.  
17-DEC-2002; 2002US-0434166P.  
24-APR-2003; 2003US-0465809P.  
(MEND-) MENDEL BIOTECHNOLOGY INC.  
Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL; Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;  
WPI; 2004-330163/30.  
P-PSDB; ADO63083.  
New recombinant polynucleotide encoding transcription factor polypeptides, useful for producing transgenic plants with advantageous properties compared to a reference plant.  
Claim 1; SEQ ID NO 1549; 510pp; English.  
The present invention relates to novel plant transcription factor proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The sequences can be used to produce transgenic plants, which overexpress (II), where the transgenic plant has an altered trait as compared to a non-transgenic plant or wild-type plant. The transgenic plant comprises an altered trait selected from increased tolerance to abiotic stress, increased tolerance to osmotic stress, increased tolerance to cold, increased germination in cold, increased tolerance to heat, increased germination in heat, increased tolerance to freezing conditions, increased tolerance to low nitrogen conditions, increased tolerance to low phosphate conditions, increased tolerance to disease, including fungal disease and particularly Erysiphe, Fusarium and Botrytis, increased tolerance to multiple fungal pathogens, increased resistance to glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA, increased sensitivity to ACC, altered sugar sensing, increased tolerance to sugars, altered carbon/nitrogen sensing, early flowering, late flowering, altered flower structure, loss of flower determinacy, reduced fertility, altered shoot meristem development, altered branching pattern, altered stem morphology, altered vascular tissue structure, reduced apical dominance, altered trichome density, altered trichome development, altered trichome structure, altered root development, altered shade avoidance, altered seed development, altered seed ripening, altered seed germination, slow growth, fast growth, altered cell differentiation, altered cell proliferation, altered cell expansion, altered phase change, altered senescence, abnormal embryo development, altered programmed cell death, lethality when overexpressed, altered necrosis patterns, increased plant size, increased biomass, large seedlings, dwarfed plants, dark green leaves, change in leaf shape, increased leaf size and mass, light green or gray leaves, glossy leaves, altered abaxial/adaxial polarity, altered seed coloration, altered seed size, altered seed shape, large seed, increased leaf wax, increased leaf fatty acids, altered seed oil content, altered seed protein content, altered seedprenyl content, altered leaf prenyl lipid content, increased anthocyanin levels, and decreased anthocyanin levels. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 1158 BP; 374 A; 261 C; 253 G; 270 T; 0 U; 0 Other;



```
Db 220 GGTGATGATGCTGTTCTCAAGGAGCTTCTGGTGTAAAGAGGAGGACGCGT 279
QY 57 AlaAlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIleValCysGlu 76
Db 280 GCGCTGGTCCAGATAGACTGGAAGAGGACTACGTCAATTTAGTATGAAAGTTTGAA 339
QY 77 LysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGlu 96
Db 340 AAGTGAAGCAAGAGGAGCACTTACAATGAGGTTTCAGACGAGCTTGTGCTGAA 399
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 400 TTTCGCACTTCCAAATAACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 459
QY 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
Db 460 ATAAGACGAGAGTATGATGCTTTAAACGCTCTCATGCTATGATGATATATATCCAAG 519
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 520 GATAAAAGAAATTCATGAGAGGCTCTCTCGGCAAGCTTAAGCGACATTAAGAA 579
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGln 171
Db 580 TTAAGAACCAACGACTCTCACTTAGGAACAGAAATTCAGAAAGAACTGCATATCCCA 639
QY 172 GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
Db 640 GAACCTGGAAGAAATTCAGTTCAGGCTTCAGAACTGATACAGAGAATGAGCACTTAT 699
QY 192 SerSerAlaGluAsnValAsn--GlyIleArgLeuProPheValLeuValLysThrSer 210
Db 700 AGCTCAGAAATGCTCCAGTGCAGTGTGTCTTCTTCTTATCTTCTTCCAGACTCGT 759
QY 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPhe 230
Db 760 CCTCAGCAACAGTACAGAGTGGAGATATCAGAAAGATATGCGCTCGTGCATTTTGATT 819
QY 231 AsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 820 AACAGCACTCCATTTGAGTCCACGACGACAATTTTGCTCCTCAAGACTATGAAG 873

RESULT 5
AAC39795
ID AAC39795 standard; DNA; 1618 BP.
XX
AC AAC39795;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 25934.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
```



Qy 211 ArglysalargValGluilleGluileSerAspSerLysPheAlaHisPheGluPhe 230  
 Db 1008 CCTCACCAACAGTAGAAGTGGAGATATCAGAAGATGTCGTCGTCATTTTGATTC 1067  
 Qy 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerLleLeuGluGlyValArg 248  
 Db 1068 AACAGCACTCAATTGAGTCCACGACGACAATTTTGTCTCAAGACTATGAAG 1121

## RESULT 6

AD062852  
 ID AD062852 standard; DNA; 1443 BP.  
 XX  
 AC AD062852;  
 XX  
 DT  
 XX  
 DE  
 XX

15-JUL-2004 (first entry)

Transcription factor G2981/2982 orthologous sequence, SEQ ID 1319.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
 KW osmotic stress tolerance; cold tolerance; heat tolerance;  
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;  
 KW glyphosate resistance; flowering; fertility; seed development; ds.  
 XX Glycine max.

WO2004031349-A2.

15-APR-2004.

18-SEP-2003; 2003WO-US030292.

18-SEP-2002; 2002US-0411837P.

17-DEC-2002; 2002US-0434166P.

24-APR-2003; 2003US-0465809P.

(MEND-) MENDEL BIOTECHNOLOGY INC.

XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;  
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;  
 XX WPI; 2004-330163/30.

XX New recombinant polynucleotide encoding transcription factor  
 PT polypeptides, useful for producing transgenic plants with advantageous  
 PT properties compared to a reference plant.

XX Claim 1; SEQ ID NO 1319; 510pp; English.

XX The present invention relates to novel plant transcription factor  
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The  
 CC sequences can be used to produce transgenic plants, which overexpress  
 CC (iii), where the transgenic plant has an altered trait as compared to a  
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises  
 CC an altered trait selected from increased tolerance to abiotic stress,  
 CC increased tolerance to osmotic stress, increased tolerance to cold,  
 CC increased germination in cold, increased tolerance to heat, increased  
 CC germination in heat, increased tolerance to freezing conditions,  
 CC increased tolerance to low nitrogen conditions, increased tolerance to  
 CC low phosphate conditions, increased tolerance to disease, including  
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
 CC increased tolerance to multiple fungal pathogens, increased resistance to  
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance  
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late  
 CC flowering, altered flower structure, loss of flower determinacy, reduced  
 CC fertility, altered shoot meristem development, altered branching pattern,  
 CC altered stem morphology, altered vascular tissue structure, reduced  
 CC apical dominance, altered trichome density, altered trichome development,  
 CC altered trichome structure, altered root development, altered shade  
 CC avoidance, altered seed development, altered seed ripening, altered seed  
 CC germination, slow growth, fast growth, altered cell differentiation,  
 CC altered cell proliferation, altered cell expansion, altered phase change,  
 CC altered senescence, abnormal embryo development, altered programmed cell

CC death, lethality when overexpressed, altered necrosis patterns, increased  
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark  
 CC green leaves, change in leaf shape, increased leaf size and mass, light  
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,  
 CC altered seed coloration, altered seed size, altered seed shape, large  
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil  
 CC content, altered seed protein content, altered seedprenyl content,  
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and  
 CC decreased anthocyanin levels. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 1443 BP; 419 A; 273 C; 343 G; 403 T; 0 U; 5 Other;

## Alignment Scores:

Pred. No.: 1.48e-38 Length: 1443  
 Score: 482.50 Matches: 106  
 Percent Similarity: 60.91% Conservative: 42  
 Best Local Similarity: 43.62% Mismatches: 70  
 Query Match: 36.58% Indels: 25  
 DB: 12 Gaps: 4

US-10-088-830-2 (1-261) x AD062852 (1-1443)

Qy 27 SerSerValProPro----- 31

Db 168 AACAGCTTCCACCTCCCTCGATATAAACTATTCTCAATTGAACCATCTCGAGTACAT 227

Qy 32 LeuProGluAlaGlyGlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAsp 51

Db 228 GCTGATGATGCTGGATCACAGCT-----TCTCTTCTGGCAGAGAAAAG 275

Qy 52 ArgLysLysGluLysAlaAlaProArgLleThrGlyTrpGlyLeuArgGluTyrSer 71

Db 276 AGAGGCGGTCAACGGGCTGTGGACCTGATAGAGTGAAGAGGTCTCCGCCAATTAGT 335

Qy 72 LysLleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluAlaAsp 91

Db 336 ATGAAAGTGTGTGAGAAAGGTAGAGAGGGGGAAGAACACACATACATAGAGTGGCGGAT 395

Qy 92 GluIleTyrSerGluLeuLysSerMetAlaHisIle-----GlyGlnGly 106

Db 396 GAGCTTGTGTGTAATTTCTGAACCAAGCATAGTGAATTGCCCTGTATCAGCAACA 455

Qy 107 PheAspGluLysAsnLleArgArgValTyrAspAlaPheAsnValLeuAlaLeu 126

Db 456 TATGATGAAAAAACATCCGCCAGGGTCTATGATGCTCTGAACGTTCTCATGGCAATG 515

Qy 127 ArgValIleAlaLysGluLysLysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyr 146

Db 516 GATATTATTTCACAGGACAAAGAAATTCATATGGAGGGGCTTCTCTGTACTACTGTG 575

Qy 147 GluLysIleLysLysLeuGluValArgLysGluValAsnLysLleArgAsnLys 166

Db 576 ATGATATTGAAGAACTAAAGCAGAGCGGCTTGGGCTCAGGAAATGAAATTGAAAGAAA 635

Qy 167 LysAlaLeuLeuGlnGluLleGluLysGlnPheAspAspLeuGlnAsnLleLysLeuArg 186

Db 636 ACAGCCTATCTCAGGAGCTTGGAGAGCAATTCGTAGGTCTTCAGAACCTTATTCAACGA 695

Qy 187 AsnGlnThrLeuGluSerSerAlaGluAsnValAsn---GlyLleArgLeuProPheVal 205

Db 696 AATGACAGTATTATAGCTTCAGGAAATCTCCAGGAGGGTGTATCTTTACCCCTTTATT 755

Qy 206 LeuValLysThrSerArgLysAlaArgValGluIleGluLleSerAspAspSerLysPhe 225

Db 756 TTGGTACAGACAGCTCCTCATGCACTGTGGAAGTGAAGTGAATATCAGAAAGATATGACGCTT 815

Qy 226 AlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerLleLeuGlu 245

Db 816 GTTCATTTTGATTTCAATAGCAGCTCTCTTTTGTAGCTGCATGACGACAAATATGTCTCAAG 875









DR P-PSDB; AAU72514.

XX New cell cycle protein and nucleic acid molecule encoding it useful for

PT regulating cell cycle progression in plants and for identifying

PT modulators which are useful as herbicides or plant growth regulators.

XX

PS Claim 38; Page 185-186; 316pp; English.

XX

CC The invention relates to a novel cell cycle protein (CCP) and the

CC polynucleotides encoding them. CCP is useful for identifying a compound

CC which modulates the activity of the polypeptide and which binds to the

CC polypeptide and an anti-CCP antibody is useful for detecting the presence

CC of CCP in a sample. A CCP modulator is useful for modulating the cell

CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,

CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP

CC nucleic acid and polypeptide molecules are useful as modulating agents in

CC regulating cell cycle progression in plants. CCP is useful to treat

CC disorders characterised by insufficient or excessive production of CCP

CC protein or production of CCP protein forms which have decreased or

CC aberrant activity. Compounds that bind to or modulate the activity of CCP

CC polypeptide are useful as herbicides or plant growth regulators. The

CC polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,

CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,

CC stimulation or enhancement of cell division, DNA replication, seed set,

CC seed size, seed development, tuber, fruit, leaf formation, shoot and root

CC initiation and/or development, module function, dwarfism in plants,

CC senescence, tolerance or resistance to stress. CCP, the polynucleotide

CC and the anti-CCP antibody are useful in agriculture to modulate the

CC protein levels or activity of a protein involved in the cell cycle due to

CC environmental conditions, including abiotic stress such as cold, nutrient

CC deprivation, heat, drought, salt stress, or biotic stress such as

CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate

CC plant architecture, plant quality traits, plant reproduction and seed

CC development, endoreduplication in storage cells, storage tissues and/or

CC storage organs of plants or its parts. CCP is useful as an immunogen to

CC generate antibodies. CCP protein is useful to screen for naturally

CC occurring CCP substrates. The polynucleotide is useful for expressing CCP

CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to

CC modulate CCP activity. The present sequence encodes a CCP protein of the

XX invention

XX

SQ Sequence 1131 BP; 366 A; 256 C; 247 G; 262 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.36e-37	Length:	1131
Score:	466.50	Matches:	101
Percent Similarity:	63.76%	Conservative:	38
Best Local Similarity:	46.33%	Mismatches:	64
Query Match:	35.37%	Indels:	15
DB:	6	Gaps:	3

US-10-088-830-2 (1-261) x AAS96304 (1-1131)

QY 37 GlyAenAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLys 56

DB 220 GGTGATGATCTGGTCTCAGAGGCTTCGGTTTAAAGAGAGAGAGGGACAGCGT 279

QY 57 AlaAlaAlaProArgLethrGlyLysGlyLeuArgLysSerLysLysValCysGlu 76

DB 280 GCGGCTGGTCCAGATAAGACTGGAAGAGGACTACGTCATTTAGTAAAGTTTGTGAA 339

QY 77 LysValGluAlaLysGlyArgThrThrThrArgGluValAlaAspGluLysSerGlu 96

DB 340 AAGGTGGAAGCAAGAGAGGACCACTTCAATGAGGTTCACACGAGCTTGTGTGAA 399

QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111

DB 400 TTTCACCTCCCAATACGATGGAACATCCCTTCATCAGCAACAGTATGATGAGAAAC 459

QY 112 IleArgArgValThrAspAlaPheAsnValLeuAlaLeuArgValIleAlaLys 131

DB 460 ATAAGACGAAGATATGATGCTTTAAACGCTCATGCTGATGATATATATATCAAG 519

QY 132 GluLysLysGluLeuArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151

DB 520 GATAAAGAAGAAATCAATGGAGAGTCTTCTCGACAGCTTAAGCCACATTGAAGAA 579

QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlu 171

DB 580 TTAAGACGAACGACTCTCCTTAGGAACAGAAATGAGAGAAATGCTATATCCCAA 639

QY 172 GluLysGluLysGlnPheAspAspLeuGlnAsnLysLeuArgAsnGlnThrLeuGlu 191

DB 640 GAACCTGGAGAACAA-----AGAAATGAGCATTATAT 672

QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210

DB 673 AGCTCAGGAATGCTCCAGTGGCGGTGTCTCTCTTTATCTCTTGTCCAGACTCGT 732

QY 211 ArgLysAlaArgValGluLeuLeuSerAspSerLysPheAlaHisPheGluPhe 230

DB 733 CCTCAGCAACAGTAGAGTGGAGATATCAGAAGATATGAGCTCGTGCATTTGATTTC 792

QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248

DB 793 AACAGCACTCCATTGTGAGCTCCAGCAGCAAAATTTGTCTCAAGACTATGAAG 846

RESULT 10

ADQ36854

ID ADQ36854 standard; DNA; 1140 BP.

XX

AC ADQ36854;

DT 07-OCT-2004 (first entry)

XX

DE Cell proliferation-related nucleic acid sequence #7.

XX

KW cell proliferation related polypeptide; cell proliferation; senescence;

KW differentiation; stress response; ds.

XX

OS Oryza sativa.

XX

PN WO2004061122-A2.

XX

PD 22-JUL-2004.

XX

PF 23-DEC-2003; 2003WO-US041200.

XX

PR 26-DEC-2002; 2002US-0436565P.

XX

PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX

PI Cooper B;

XX

DR WPI; 2004-534388/51.

XX

PT New nucleic acid molecule encoding a cell proliferation-related

PT polypeptide, useful for modulating cell proliferation, senescence,

PT differentiation, development, and stress response in plants, and for

PT producing enhanced food crops.

XX

PS Claim 3; SEQ ID NO 13; 408pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a cell proliferation-related polypeptide. The nucleic acid

CC molecule and the encoded polypeptide, and methods are useful for

CC modulating cell proliferation, senescence, differentiation, development,

CC and stress response in plants, and for producing enhanced food crops. The

CC present sequence represents a cell proliferation-related nucleic acid

CC sequence. The present sequence is published separately from the main body

CC of the specification as EPO data.

XX

SQ Sequence 1140 BP; 327 A; 268 C; 310 G; 235 T; 0 U; 0 Other;

Alignment Scores:

```
Pred. No.: 4,41e-37 Length: 1140
Score: 466.50 Matches: 111
Percent Similarity: 52.38% Conservative: 43
Best Local Similarity: 37.76% Mismatches: 77
Query Match: 35.37% Indels: 63
DB: 12 Gaps: 8

US-10-088-830-2 (1-261) x ADQ36854 (1-1140)
QY 3 ProProArgGlyGlyAlaAlaAlaAlaThraAlaAlaLeuAspLeuThrGlyValHis 22
Db 109 CCCTCTCGGCGGAGCGCAATCGCTTCACGAGCGG-----147
QY 23 IleLeuGluAlaSerSerValProPro-----LeuProGluAla 35
Db 148 -----GGGAGCGGCGCTCCCGTCCAGCGCGAGCAGCATGTCCCGCAGCC 198
QY 36 GlyGlyAsnAlaValGlnArgLysGlyAlaValAspPro-----48
Db 199 GCAGGCATGCGCGGCGGCGGCGGCGCTCTACTCCGATTAGTGAATACCTTCTC 258
QY 48 -----48
Db 259 CGCTCAACGACCTTGACATCCAGCGGCGAGCGCTTCTCAGCGCTCCACGAGC 318
QY 49 AspLysAspArgLysLysGluLysAlaAlaAlaProArgIleThrGlyTrpGlyLeuArg 68
Db 319 AAGAAGAAGAGAGAGGAGCAGCAGCAGTGTGCTTCACAAAGGTGCGAGGGCGCTGCG 378
QY 69 GluTyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGlu 88
Db 379 CAGTTAGTATGAAGTTTGTGAGAAGTTGAAAGTAAAGGGAGAAACAACATACGAG 438
QY 89 ValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHisIle-----103
Db 439 GTGCGAGATGAACCTGTGTCGGAATTTGCGAGATCCCAATACAGCATTTTGCCACCAGAT 498
QY 104 -----GlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAsp 118
Db 499 CCGGATAATCCCAATGCACACATATGACGAGAAAATATACGAGAGAGGTTTATGAT 558
QY 119 AlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLysGluIleArgTrp 138
Db 559 GCTCTGAATGTTCTGATGCTATGGAGATTATATCTAAAGATAAAAGAAATTCAGTGG 618
QY 139 MetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluValArgLysGlu 158
Db 619 AAGGGGTGCTCGCA-----ACCAGTATAAATGATATTGAAGATTGCGACGGA 669
QY 159 LeuVal-----AsnLysIleArgAsnLysAlaLeuLeuGlnGluLys 175
Db 670 CTGTAGGACTGAAAGTAGGATTGAGAGAAAATACATATTGCGAGCTGCAAGAC 729
QY 176 GlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGlu 195
Db 730 CAATTGTAGGTATGCAAAAGTTGATACAAAGAAATGAACAGCTATATGTTTCAGCA---786
QY 196 AsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSerArgLysAla 213
Db 787 AACATTCTCGGCTGGAGTTGCATTACCATTTATCTTCTGTTTCAGACACGGCTCATGCA 846
QY 214 ArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPheAsnGlyAla 233
Db 847 ACTGTGGAAGTTGAATATACAGAGATATGCACTTGATCACTTTGACATTTAATACACA 906
QY 234 ProPheThrLeuHisAspLeuSerIleLeuGluGlyVal 247
Db 907 CCATTGTGATGATGACTCATTTGTACTGAAAGCAATG 948

RESULT 11
AD063363
ID AD063363 standard; DNA; 1245 BP.
XX
```

```
AC AD063363;
XX 15-JUL-2004 (first entry)
XX Transcription factor G2981 orthologous sequence, SEQ ID 1830.
DE Plant; transcription factor; transgenic plant; abiotic stress tolerance;
XX osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
XX glyphosate resistance; flowering; fertility; seed development; ds.
XX Zea mays.
OS WO2004031349-A2.
PN 15-APR-2004.
PD 18-SEP-2003; 2003WO-US030292.
XX PF 18-SEP-2002; 2002US-0411837P.
XX PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann J, Haake V, Dubell AN, Keddle JS, Sherman BK,
XX WPI; 2004-330163/30.
DR New recombinant polynucleotide encoding transcription factor
XX polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
PT Claim 1; SEQ ID NO 1830; 510pp; English.
XX The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
```

XX SQ Sequence 1245 BP; 414 A; 225 C; 308 G; 298 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,11e-36 Length: 1245  
Score: 463.00 Matches: 112  
Percent Similarity: 54.80% Conservative: 42  
Best Local Similarity: 39.86% Mismatches: 79  
Query Match: 35.10% Indels: 48  
DB: 12 Gaps: 8

US-10-088-830-2 (1-261) x ADO63363 (1-1245)

QY 5 ArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGlyValHisIleLeu 24  
DB 4 CGAGGGCAGACGACGCTTCCTCCGCTTGAATATTCGACATCAACGCG----- 51  
QY 25 GluAlaSerValProLeuProGluAlaGlyGlyAsnAlaValGlnArgLysGly 44  
DB 52 -----GACGACGCGCGTCGTCGAGCTCTACGAGCAAGAAAGAGGAGGC 102  
QY 45 -----AlaValAspProAspLysAspArgLysLysGlyLysAlaAlaProArgIle 62  
DB 103 ACACGGCGAGTGGTCTCGATAAGGTAAACCGG----- 135  
QY 63 ThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAlaLysGly 82  
DB 136 -----GGACTGCGCCAGTTTGTATGTAAGTTTGTGAGAAGTTGAAAGTAAAGG 186  
QY 83 ArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHis 102  
DB 187 AGAACACATATATAGGTGGCAGATGAATCTGTGTGAGTTTACAGACCCCAACAT 246  
QY 103 -----IleGlyGlnGlyPheAspGluLysAsnIle 112  
DB 247 AATATTGAGCCACGATCCTGATACCTTAACCGCAACATATGATGAGAAATATA 306  
QY 113 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGlu 132  
DB 307 CGACGAGAGTTTATGATGCTTTCGATGTTTGTGATGATGATATATATCAAGAT 366  
QY 133 LysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeu 152  
DB 367 AAAAGGAGATCCAGTGGAGGAGGCTTGCCTG-----ACTAGTATAGTGACAT 417  
QY 153 GluGluValArgLysGluLeuVal-----AsnLysIleArgAsnLysLysAlaLeu 169  
DB 418 GAAGAAATGAAGACGGAGCTTGTGGGACTGAAAGTAGGATGAAAGAAAGTGTAT 477  
QY 170 LeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn---Gln 188  
DB 478 CTACAGGAGGTACAAAGCAATATGTAGGTCTGCAAAACCTGATCAACGAAATGAAC 537  
QY 189 ThrLeuGluSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLys 208  
DB 538 TCATATGTTTCAGAAACACGCTTCTGTGGTGGAGTGGCTTGCATTTATCCTAGTTCAG 597  
QY 209 ThrSerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHisPhe 228  
DB 598 ACCGACCTCATGCTACCTGGAAGTGTGATATCAGAGATATGACAGTGGTGCATTTT 657  
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248  
DB 658 GACTTCAATAGCACCCCATTCGAGCTCGACGACATCATACGTCCTTAAAGAAATGCGA 717  
QY 249 -----ArgAsnSerIleGlyArgAlaGlyArgAla 258  
DB 718 TTCTGTGAAGAACAAACATGACGACCTCAAGAGTCGATATCAATGAGGTGAGAGC 777  
QY 259 Thr 259  
DB 778 TCA 780

## RESULT 12

AAA59702  
ID AAA59702 standard; DNA; 900 BP.

XX  
AC AAA59702;  
DT 14-NOV-2000 (first entry)  
XX  
DE DNA encoding a transcription factor designated DP.

XX DP polypeptide; transcription factor; gene transcription; cell cycle;  
KW DNA replication; DRTF; differentiation-regulated transcription factor 1;  
KW transgenic plant; transformation efficiency; ss.  
XX  
OS Zea mays.

XX  
FH Key Location/Qualifiers  
CDS 1..900  
FT /tag= a  
FT /transl\_except= (pos: 1..3, aa: Xaa)  
FT /product= "DP polypeptide"  
FT /note= "Xaa is any amino acid"

XX  
PN WO200047614-A1.

XX  
PD 17-AUG-2000.

XX  
PF 11-FEB-2000; 2000WO-US003651.

XX  
PR 12-FEB-1999; 99US-0119857P.

XX  
PA (PION-) PIONEER HI-BRED INT INC.

XX  
PI Lowe KS, Gordon-Kamm WJ, Bailey MA, Gregory CA;

XX  
DR WPI; 2000-524530/47.  
DR P-PSDB; AAB07975.

XX Transgenic plants comprising an expression cassette consisting of a DP  
PT nucleic acid, when expressed, the nucleic acid increases the efficiency  
PT of plant cell transformation by increasing the number of dividing cells  
PT in the plant.

XX  
PS Claim 6; Page 70-71; 76pp; English.

XX The present sequence encodes a corn DP (not defined) polypeptide. The  
CC polypeptide is a transcription factor that activates transcription of  
CC numerous genes involved in DNA replication, thus playing a role in the G1  
CC to S transition in the cell cycle. Do, also called DRTF (differentiation-  
CC regulated transcription factor 1) has been shown to form specific  
CC multiprotein complexes with the retinoblastoma susceptibility protein,  
CC p107, cyclins and cdk2. The DP polynucleotide was used to produce a  
CC transgenic plant. The DP nucleic acids and proteins are useful for  
CC increasing transformation efficiency in plants by increasing the number  
CC of dividing cells in the plant (dividing cells may be more receptive to  
CC transformation)

XX  
SQ Sequence 900 BP; 293 A; 178 C; 231 G; 195 T; 0 U; 3 Other;

## Alignment Scores:

Pred. No.: 1.16e-36 Length: 900  
Score: 461.00 Matches: 110  
Percent Similarity: 55.93% Conservative: 41  
Best Local Similarity: 40.74% Mismatches: 71  
Query Match: 34.95% Indels: 48  
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x AAA59702 (1-900)

QY 16 LeuAspLeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAla 35  
DB 31 CTCGACATCAACGCG-----GACGACGCGCGCTGTCGACGCT 69







Db 444 GGTAGATTGAGAGAAAGCTGTTACTTACAGAGCTACAAGATCAATATGTAGTTTG 503  
 QY 181 GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsn---Gly 199  
 Db 504 CAAACTGATTCAAGAAATGAGCAATATATGTTTCAGGAACACACCTCTCTGGTGA 563  
 QY 200 IleArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIle 219  
 Db 564 GTGGTTTGGCAATCATCTCTAGTCCAGACCCGACCTCATCTACCTGGGAAGTTGAGATA 623  
 QY 220 SerAspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAsp 239  
 Db 624 TCAGAGATATGACGTGGTGCATTTGACTTCAATAGCACCCCATTCGAGCTGCACGAC 683  
 QY 240 AspLeuSerIleLeuGluGlyValArg-----Arg 249  
 Db 684 GATCATAGCTCTAAAGAAATGCGATTTCTGTGGAAGACAAACATGACAGACTCAA 743  
 QY 250 AsnSerIleGlyArgAlaGlyAlaThr 259  
 Db 744 GAGTCGATATCAAAATGAGGTGAGAGCTCA 773

## RESULT 15

AAS96287  
 ID AAS96287 standard; cDNA; 1114 BP.

AC AAS96287;

DT 26-FEB-2002 (first entry)

XX Arabidopsis cDNA encoding partial cell cycle protein CCP16.

XX Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;  
 KW growth regulator; plant development; abiotic stress; biotic stress;  
 KW nutrient deprivation; pathogen attack; crop yield.

XX Arabidopsis thaliana.

XX WO200185946-A2.

PD 15-NOV-2001.

XX 14-MAY-2001; 2001WO-1B001307.

PF 12-MAY-2000; 2000US-0204045P.

XX (CROP-) CROPEDESIGN NV.

XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX WPI; 2002-062249/08.

DR P-PSDB; AAU72497.

XX New cell cycle protein and nucleic acid molecule encoding it useful for  
 PT regulating cell cycle progression in plants and for identifying  
 PT modulators which are useful as herbicides or plant growth regulators.

XX Claim 38; Fig 16; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the  
 CC polynucleotides encoding them. CCP is useful for identifying a compound  
 CC which modulates the activity of the polypeptide and which binds to the  
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP...  
 CC nucleic acid and polypeptide molecules are useful as modulating agents in  
 CC regulating cell cycle progression in plants. CCP is useful to treat  
 CC disorders characterised by insufficient or excessive production of CCP  
 CC protein or production of CCP protein forms which have decreased or  
 CC aberrant activity. Compounds that bind to or modulate the activity of CCP  
 CC polypeptide are useful as herbicides or plant growth regulators. The  
 CC polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, module function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due to  
 CC environmental conditions, including abiotic stress such as cold, nutrient  
 CC deprivation, heat, drought, salt stress, or biotic stress such as  
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate  
 CC plant architecture, plant quality traits, plant reproduction and seed  
 CC development, endoreduplication in storage cells, storage tissues and/or  
 CC storage organs of plants or its parts. CCP is useful as an immunogen to  
 CC generate antibodies. CCP protein is useful to screen for naturally  
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP  
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to  
 CC modulate CCP activity. The present sequence encodes a CCP protein of the  
 CC invention  
 XX  
 SQ Sequence 1114 BP; 378 A; 197 C; 243 G; 296 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2.77e-33 Length: 1114  
 Score: 428.50 Matches: 88  
 Percent Similarity: 64.32% Conservative: 49  
 Best Local Similarity: 41.31% Mismatches: 69  
 Query Match: 32.49% Indels: 7  
 DB: 6 Gaps: 2

US-10-088-830-2 (1-261) x AAS96287 (1-1114)

QY 40 ValGlnArgLysGlyAlaValAlaAspProAspLysAspArgLysLysGluLysAlaAlaA 59  
 Db 92 GTGAGAGGAAATGATTTGATGATGATTCGAAATTTGGATCGAGAGAAAGGCCAA 151  
 QY 60 ProArgIleThrGlyTrpGlyLeuArgGluTrpSerLysIleValCysGluLysValGlu 79  
 Db 152 TCAGAGAACTTCGGAGCGGGCTTCGTCATTCAGTTGTTGTTGTCAGAGTTGGAA 211  
 QY 80 AlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99  
 Db 212 GCCAAGAAGATAACTACTTACAGAGGAGTTGCAGACGAAATTTATTCAGATTTTGCACA 271  
 QY 100 MetAlaHisIleGlyCln-----GlyPheAspGluLysAsnIleArg 113  
 Db 272 ATTAAGCAAAACGCGAGAGAGCCTTTGAATGAAATGAGTACAAATGAGAGAACATAAGG 331  
 QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133  
 Db 332 CGGAGAGTCTACGATGCGCTCAATGTGTTTCATCGCGCTTGATATTTTCAAGGGATAA 391  
 QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153  
 Db 392 AAGGAATCCGGTGGAAAGACTTCTTATTCCTGCAAAAAGGATGTGGAAGAAGTCAAG 451  
 QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle 173  
 Db 452 ATGGATCGGTAATAAGTTATGAGCAGTGTGCAAAAGAGCGCTTTCTTAAAGAGTTG 511  
 QY 174 GluLysGlnPheAspAspLeuGlnAsnIleLysLysLeuArgAsnGlnThrLeuGluSerSer 193  
 Db 512 AGAGAAAAGGTCTCAAGTCTTCAGAGCTTATCTCGAGAAATCAAGAGATGGTTGTGAAG 571  
 QY 194 AlaGluAsn---ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212  
 Db 572 ACTCAAGGCCGACGAGAGGATTTACCTTACCATTCATCTACTTGAGACAAACCTCAC 631  
 QY 213 AlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPheAsnGly 232  
 Db 632 GCAGTAGTCGAATCGAGATTTCTGAAGATATGCAACTTTGTACACCTCGACTTCAATAGC 691  
 QY 233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245

Db  
692 ACACCTTCTCGGTC CATGATGATGCTTACATTTGAAA 730

Search completed: February 6, 2005, 22:04:22  
Job time : 588 secs



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	844.5	64.0	1119	9	CL959517	OsIFCC003
2	817	61.9	649	6	CD882937	Fl.111102
3	694	52.6	612	7	CR289243	CR289243
4	651	49.4	765	6	CA764996	AF53-Rbf
5	615.5	46.7	607	2	BE456002	HVSME001
6	611.5	46.4	665	6	CA078334	SCRLAM100
7	575.5	43.6	548	2	BES11893	BE511893
8	567	43.0	634	7	CF484481	POLL_25_C
9	505.5	38.3	680	6	CA790160	SCAGFL801



```

Db 122 AAAGAAAAAGAGATACGGTGGATGGGCTTTCAAAATTACAGATATGAAAAATAAG 181
QY 151 LysLeuGluValArgLysGluLeuValAsnLysLeuValAsnLysLeuValAsnLysLeu 170
Db 182 AAGCTTGAGGAGTTGTAAGAACATGTCACAGATTAGGAACAGAGGACCTCTC 241
QY 171 GlnGluLeuLysGlnPheAspAspLeuGlnAsnLysLeuValAsnGlnThrLeu 190
Db 242 CAGGAATCGAAAAACAGTTGGATGATCTCCAAAACATCATGTTACGTAACCAACACTG 301
QY 191 GluSerSerAlaGluAsnValAsnGlnLysLeuValAsnGlnThrLeu 210
Db 302 GAAAGCTCAGCAGAGATGTAATGGATCGATCGGCTTCCATTCGATGGGTCAAGACATCT 361
QY 211 ArgLysAlaArgValGluLeuGluLeuSerAspAspSerLysPheAlaHisPheGluPhe 230
Db 362 AGGAAGCAGAGGTGGAATGAGATTTCAGACATCAAGTTGCCCATTTCCGAGTTTC 421
QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerLysLeuGluValArgAsn 250
Db 422 AATGGTGCACCATTCACATTGCATGATGATCTCTCAATCTTCGAGGGGATAAGGCGTAAC 481
QY 251 SerLysGlyArgAlaGlyArgAlaThrLeuHis 261
Db 482 AGCATAGGAAGAGCTGGCGCGCCACCTTCAC 514

```

```

RESULT 3
LOCUS CR289243 612 bp mRNA linear EST 27-FEB-2004
DEFINITION CR289243 Oryza sativa library (Han B) Oryza sativa cDNA clone
P704a10p5, mRNA sequence.
ACCESSION CR289243
VERSION CR289243.1 GI:44675809
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa

```

```

REFERENCE 1 (bases 1 to 612)
AUTHORS Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J.,
Zhao, Q. J., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L.,
Weng, Q. J., Zhang, L., Lu, X. Q., Yu, S. L., Liu, X. H., Lu, T. T.,
Zhang, Y. J., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X.,
Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.

```

```

TITLE Rice cDNA EST clone
COMMENT Unpublished (2003)
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhanncgr.ac.cn
Clone requests: bhanncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.

```

```

FEATURES
source 1. .612
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="p704a10p5"
/location/Qualifiers
/location_lib="Oryza sativa library (Han B)"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2,37e-67 Length: 612
Score: 694.00 Matches: 134
Percent Similarity: 87.91% Conservative: 26
Best Local Similarity: 73.63% Mismatches: 22
Query Match: 52.62% Indels: 0
DB: 7 Gaps: 0

```

```

US-10-088-830-2 (1-261) x CR289243 (1-612)
QY 70 TyrSerLysLeuValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89
Db 2 TTCAGCAAGATAGTTTCTAAGAAAGTTGAGGCCAAGAGGAACACACATATATAGGTT 61
QY 90 AlaAspGluLeuTyrSerGluLeuLysSerMetAlaHisLysGlnGlyPheAspGlu 109
Db 62 GCCGTAAGATTTTGGCGGAGCTGAAGTCCATTACGCAGAACGGCTGGAGTTTGATGAG 121
QY 110 LysAsnLeuArgArgValTyrAspAlaPheAsnValLeuLeuAlaLeuArgValLe 129
Db 122 AGAAATATTAGCGGGGGGTATATGATGCTTTCAATGTGCTCATTCGAATTCGTGTATT 181
QY 130 AlaLysGluLysLysGluLeuArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLe 149
Db 182 GCAAAAGATATAAAGAGATAAAGTGGATGGGCTTACTTAATATATAGATACGAAGAATA 241
QY 150 LysLysLeuGluGluValArgLysGluLeuValAsnLysLeuValAsnLysLysAlaLeu 169
Db 242 CAGAAGTTGGAGGAGTTTCAAAAGAACTCATCCAGGATCAAGTAATAAGAAAGAGCTT 301
QY 170 LeuGlnGluLeuGluLysGlnPheAspAspLeuGlnAsnLysLeuArgAsnGlnThr 189
Db 302 CTCAGGAATTTGAAAAGCAGTTTGTATGACCTTCAGAAATATTACATTACGCAACAGGCT 361
QY 190 LeuGluSerSerAlaGluAsnValAsnGlyLysLeuArgLeuProPheValLeuValLysThr 209
Db 362 AGTCAGAGCCAGCAGAAAGTGTAAATGGATCTCTTCCGTACTTATTTGATCAAGACA 421
QY 210 SerArgLysAlaArgValGluLeuGluLeuSerAspAspSerLysPheAlaHisPheGlu 229
Db 422 TCCGAAAAGCAAGGGTGGAACTTGAATTCGAAAGATTCCAACTTTGCACGGTTCCGAC 481
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerLysLeuGluGlyValArg 249
Db 482 TTCAACTGGCCCAATCACTGATGATATATATCAATCTTGAATCATCATCGCT 541
QY 250 AsnSer 251
Db 542 ACCACC 547

```

```

RESULT 4
CA764996
LOCUS DEFINITION

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

CA764996 765 bp mRNA linear EST 08-JAN-2003  
AF53-Rpf 05 M03 T7 008.ab1 IRRI Drought Stress Panicle Library  
Oryza sativa (indica cultivar-group) cDNA clone C0001827.5' similar  
to transcription factor DP-1 (E2F dimerization partner 1) (DR1F1-  
polypeptide-1), mRNA sequence.

CA764996  
CA764996.2 GI:27546711

EST.  
Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 765)  
Bennett, J., Arumugam, K., Lafitte, R., Wen, J., Rudd, S. and  
Bruskewich, R.M.

IRRI Drought Stress Panicle cDNA Library  
Unpublished (2002)

On Dec 2, 2002 this sequence version replaced gi:25994251.

Contact: Richard Bruskewich  
Biometrics and Bioinformatics Unit  
International Rice Research Institute  
DAPO 7777, Metro Manila, Philippines  
Tel: +63-2-845-0563  
Fax: +63-2-845-0606

Email: r.bruskewich@cgiar.org  
International Rice Information System (IRIS)  
http://www.iris.irri.org; D0201826

Assignment of putative function to the sequence by S. Rudd of the

## Munich Information Center for Protein Sequences

(http://mips.gsf.de)

Plate: 05 Row: M Column: 03.

Location/Qualifiers

## FEATURES

source

1. 765  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="IR64"  
 /db\_xref="taxon:39946"  
 /clone="C0001827"  
 /tissue\_type="Panicles"  
 /dev\_stage="Flowering"  
 /clone\_lib="IRRI Drought Stress Panicle Library"  
 /note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,15e-62 Length: 765  
 Score: 651.00 Matches: 125  
 Percent Similarity: 89.57% Conservatives: 21  
 Best Local Similarity: 76.69% Mismatches: 17  
 Query Match: 49.36% Indels: 0  
 DB: 6 Gaps: 0

US-10-088-830-2 (1-261) x CA764996 (1-765)

QY 94 TyrSerLeuLysSerMetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113  
 DB 3 TTTGGGAGCTGAAGTCCATTCGCGAACGGCTGGAGTTTGATGAGAAATATTAGG 62  
 QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133  
 DB 63 CGGAGGGTATATGATGCTTTCAATGTGCTCATTTGCAATTCGTGTTATGCAAAAGATAA 122  
 QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153  
 DB 123 AAGGAGATAAAGTGGATGGCCCTTACTAATTATAGATACGAAAGATACAGAAGTTGGAG 182  
 QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLysGlnGluIle 173  
 DB 183 GAGTTCCACAAAGACTCATCCAGGATCAGAAATAGAAAGCTTCTCCAGGAATTT 242  
 QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193  
 DB 243 GAAAGCAGTTGATGACCTTCAGATATATTACATTACGCAACAGCGCTAGTCAGAGGCCA 302  
 QY 194 AlaGluAsnValAsnGlyIleArgLeuPheValLeuValLysThrSerArgLysAla 213  
 DB 303 GCAGAAAGTGTAAATGACATCCCTTCCTTCTTATTGATCAGACATCCCGAAAGCA 362  
 QY 214 ArgValGluIleGluLysSerAspSerLysPheAlaHisPheGluPheAsnGlyVala 233  
 DB 363 AGGTGGAATTTGAGATTCGGAAGATTCGAAGTTTGACGGTTGACCTTCACGGTGCA 422  
 QY 234 ProPheThrLeuHisAspLeuSerIleLeuGluGlyValArgArgAsnSerIleGly 253  
 DB 423 CCATTCCACATGATGATGATCAATCTTTGAGCCATCAGCGGTACCAAAAGGA 482  
 QY 254 ArgAlaGly 256  
 DB 483 AGAGCTGGC 491

## RESULT 5

BE456002

## LOCUS

DEFINITION

BE456002 607 bp mRNA linear EST 22-OCT-2001  
 HVSDNA0008 (white to yellow anther) Hordeum vulgare subsp. vulgare  
 cDNA clone HVSMEG0019D17f, mRNA sequence.  
 BE456002

## ACCESSION

## VERSION

KEYWORDS

SOURCE

## ORGANISM

## REFERENCE

AUTHORS

## TITLE

## JOURNAL

COMMENT

## FEATURES

source

BE456002.3 GI:16318905

EST.

Hordeum vulgare subsp. vulgare

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 607)

Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,  
 Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,  
 Fenton,R.D., Close,S.J., Oates,R. and Main,D.

Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex pre-anthesis spike cDNA library  
 Unpublished (2001)

On Jul 26, 2000 this sequence version replaced gi:13155022.

Contact: Wing RA

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 141

Seq primer: AATTAACTCTCACTAAAGG

High quality sequence stop: 585.

Location/Qualifiers

1..607

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Morex"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="HVSMEG0019D17f"

/tissue\_type="pre-anthesis spike"

/lab\_host="SOLR"

/clone\_lib="Hordeum vulgare pre-anthesis spike EST library  
 HVCDNA0008 (white to yellow anther)"

/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
 Plants were grown in the greenhouse at the University of  
 California, Riverside (Fenton, SJ Close, TU Close). Whole  
 spike with awns trimmed were collected at white, green and  
 yellow anther stages (Fenton). Total RNA was prepared from  
 each pool, equal quantities of all three RNA pools were  
 combined, poly(A) RNA was purified from the mixture, one  
 primary unamplified cDNA library was made, and 1 million  
 pfu were in vivo excised to give pBluescript SK(-) cDNA  
 phagemids. These steps were performed in the TU Close lab  
 (Choi) at the University of California, Riverside.

Phagemids were plated and picked at the Clemson University  
 Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins  
 and Wing) Plasmid DNA preparations, DNA sequencing and  
 sequence analysis were performed at CUGI (Wing, Yu,  
 Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence  
 has been trimmed to remove vector sequence and contains a  
 minimum of 100 bases of phred value 20 or above. For more  
 details on library preparation and sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders\_Also  
 see Close TJ, Wing R, Kleinhofs A, Wise R (2001)

Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.51e-58 Length: 607  
 Score: 615.50 Matches: 137  
 Percent Similarity: 84.30% Conservatives: 8  
 Best Local Similarity: 79.65% Mismatches: 22  
 Query Match: 46.66% Indels: 5  
 DB: 2 Gaps: 1

US-10-088-830-2 (1-261) x BE456002 (1-607)













```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1418)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source
1..1418
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL81ZA03"
/tissue type="Adult vegetative tissue"
/plasmid="PCWSPORT_6"
1..1418
/gene="At5g03415"

ORIGIN
Alignment Scores:
Pred. No.: 1,68e-42 Length: 1418
Score: 477.00 Matches: 104
Percent Similarity: 65.75% Conservative: 40
Best Local Similarity: 47.49% Mismatches: 68
Query Match: 36.16% Indels: 7
DB: 3 Gaps: 2

US-10-088-830-2 (1-261) x CNS09YMC (1-1418)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
|:::|
|:::|
|:::|
Db 391 GGTGATGATGCTGTTCTCAAGAGCTTCTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 450
|:::|
|:::|
|:::|
QY 57 AlaAlaAlaProArgLysThrGlyLeuArgLysLysLysLysLysLysLysLysLysLys 76
|:::|
|:::|
|:::|
Db 451 GCGGCTGGTCCAGATGAAGACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
|:::|
|:::|
|:::|
QY 77 LysValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThr 96
|:::|
|:::|
|:::|
Db 511 AAGTGGGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
|:::|
|:::|
|:::|
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
|:::|
|:::|
|:::|
Db 571 TTGTCACCTCCAAATACGATGGAACATCCCTGATCAGCAGACAGATGATGAGAAAC 630
|:::|
|:::|
|:::|
QY 112 IleArgArgValTyrAspAlaPheAsnValLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 131
|:::|
|:::|
|:::|
Db 631 ATAAGACGAAGAGATATGATGCTTTAAACGCTCCTCATGCTATGATGATATATATCAAG 690
|:::|
|:::|
|:::|
QY 132 GluLysLysGluLeuArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLys 151
|:::|
|:::|
|:::|
Db 691 GATAAAAAAGAAATCAATGGAGAGAGCTCTCTCGGACCAAGCTTAAGCCACATTGAAGA 750
|:::|
|:::|
|:::|
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGln 171
|:::|
|:::|
|:::|
Db 751 TTAAGACGACGACATCTCTACCTAGGACAGAGATTGAGAGAAACTGCATATATCCCAA 810
|:::|
|:::|
|:::|
QY 172 GluLysGluGlnPheAspLysLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
|:::|
|:::|
|:::|
Db 811 GAACCTGGAAGAACAAATATGTAGGCTTTCAGAACTCTGATATACAGAGAAATGAGCATTAT 870
|:::|
|:::|
|:::|

QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuPro-PheValLeuValLysThrSe 210
|:::|
|:::|
|:::|
Db 871 AGCTCAGGAATGCTCCAGTGGCGGTGTCTCTTCCTTTTATCTCTGTCAGACTCG 930
|:::|
|:::|
|:::|
QY 210 rArgLysAlaArgValGluIleGluLysSerAspSerLysPheAlaHisPheGluPh 230
|:::|
|:::|
|:::|
Db 931 TCCTCAGCAACAGTAGAATGAGATATCAGAAGATATGCAGCTCGTGCATTTTGATT 990
|:::|
|:::|
|:::|
QY 230 eAsnGlyAlaProPheThrIleuHisAspLysLeuSerIleLeuGluGlyValArg 248
|:::|
|:::|
|:::|
Db 991 CAACAGCACTCCATTTGAGCTCCAGCAGCACTATTTTGTCTCAGACATATGAAG 1045
|:::|
|:::|
|:::|

RESULT 12
CNS09ZIR 1568 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTSL732B01 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX831265.1 GI:42456015
VERSION 1
KEYWORDS HTC; GSUT; cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1568)
AUTHORS Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1568)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
Location/Qualifiers
1..1568
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL732B01"
/tissue type="Adult vegetative tissue"
/plasmid="PCWSPORT_6"
1..1568
/gene="At5g03415"

ORIGIN
Alignment Scores:
Pred. No.: 1,93e-42 Length: 1568
Score: 477.00 Matches: 103
Percent Similarity: 65.30% Conservative: 40
Best Local Similarity: 47.03% Mismatches: 69
Query Match: 36.16% Indels: 7

```



us-10-088-830-2.p2n.rst

**AUTHORS**  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

**TITLE**  
Generation of ESTs from abiotic stressed potato tissue

**JOURNAL**  
Unpublished (2003)

**COMMENT**  
Other ESTs: EST724740  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr. Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1. .992  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAE738"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1 d,  
2d and 4d). Set 2 were grown under the standard conditions and  
were water stressed by withdrawal of further watering  
applications. Drought-stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

**FEATURES**  
source

## ORIGIN

Alignment Scores:	
Pred. DB: 1.77e-41	Length: 992
Score: 466.00	Matches: 100
Percent Similarity: 64.45%	Conservative: 37
Best Local Similarity: 47.39%	Mismatches: 66
Query Match: 35.33%	Indels: 8
DB: 7	Gaps: 3
US-10-088-830-2 (1-361)	X CK278661 (1-992)

QY	50	LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu	69
		:::::	:::::
Db	53	AAAAAGAAAGAGCTCAGCGTGTACTTCGGAGGGCATAGAAGTGTGTAGAGCACTCAGAC	112
		:::::	:::::
QY	70	TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal	89
		:::::	:::::
Db	113	TTCCAGCATGAAGTTTGTAAGAGTGGAAAGCAAGAGAACACTAGTATATATGAGTTT	172
		:::::	:::::
QY	90	AlaAspGluIleTyrSerGluLeuLysSerMetAlaHisIleGlyGingly-----	106
		:::::	:::::
Db	173	GCTGATGAACCTGTAGCTGAGTTTTCATGCTACCATAAGTGTGCAGGCCACATCAG	232
		:::::	:::::
QY	107	-----PheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIle	124
		:::::	:::::
Db	233	AACAATATATGAGAAGAACATCAGACGAGCTCACGATGCTCTGGAAGCTACTTATG	292
		:::::	:::::
QY	125	AlaLeuArgValIleAlaLysGluLysGluIleArgTrpMetGlyLeuSerAsnTyr	144

Db	293	GCATGATGATCATTTCTAAAGATAAAAGGAATACAGTGGAGGATTACCAAGCACT	352
Qy	145	ArgTyx-GluLysIleLysLysLeuGluValArgLysGlnLeuValAsnLysIleArg	164
Db	353	GATGCAATGATATTGAGGAGCTAAAGACTGAGCGTCTTAACCTTGAGAAATAGGATTGAA	412
Qy	165	AsnLysLysAlaLeuLeuGlnGluLysGlnPheAspLeuGlnAsnIleLys	184
Db	413	AAGAAAGACGACCTATTATTGAAGAAGAACTTTGAAGATCAATATGAGCGCTTCAAAACCTCAT	472
Qy	185	LeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsn--GlyIleArgLeuPro	203
Db	473	AAAGCAATGATCAGTTGTATGCTCAGCAATGCTCCTAGTGGTGGTGGCTTTACCG	532
Qy	204	PheValLeuValLysThrSerArgLysAlaArgValGluLysLeuSerAspAspSer	223
Db	533	TTTATTATTAGTCAGACTCGCTCATGCAACAGTTGAAGTGGAAATATCAGAAGATATG	592
Qy	224	LysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisaspLeuSerIle	243
Db	593	CAGCTGGTGATTTGCATTTCAACAGCACTCCGTTGAGCTACATGACGATAATTATATC	652
Qy	244	LeuGluGlyValArgArgAsnSerIleGlyArg	254
Db	653	CTCAAGCAATG-----AATTTTGGCGAAGA	679
RESULT	15		
LOCUS	AY108383		
DEFINITION	Zea mays PC0085007 mRNA sequence.	1245 bp	linear
ACCESSION	AY108383		
VERSION	AY108383.1	GI:21211451	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1245)		
AUTHORS	Coe, E.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.		

**FEATURES**  
**source**

## ORIGIN

Alignment Scores:

Pred. No.: 5.25e-41 Length: 1245  
Score: 463.00 Matches: 112  
Percent Similarity: 54.80% Conservative: 42  
Best Local Similarity: 39.86% Mismatches: 79  
Query Match: 35.10% Indels: 48  
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x AV108383 (1-1245)

```
Qy 5 ArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGlyValHisIleLeu 24
Db 4 CGAGGCGACGACGACGCTTCCTCCGCTTGATAATTCGACATCAACGCGC-----51

Qy 25 GluAlaSerSerValProProLeuProGluAlaGlyGlyAsnAlaValGlnArgLysGly 44
Db 52 -----GACGACGCGCGCTCGTCGCGAGCTCTACGAGCAAGAGAAAGGAGAGGC 102

Qy 45 -----AlaValAspProAspLysAspArgLysLysGlyLysAlaAlaProArgIle 62
Db 103 ACACGGCGAGTGGTCTCTGATAAGGTAAACCGG-----135

Qy 63 ThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGlyLysValGluAlaLysGly 82
Db 136 -----GGACTGCGCCAGTTTAGTATGAAAGTTTGTGAGAAAGTTGAAAGTAAAGGG 186

Qy 83 ArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHis 102
Db 187 AGAACACATATAATAGGTGGGAGATGAACCTTGTGCTGAGTTTACAGACCCCAACAAT 246

Qy 103 -----IleGlyGlnGlyPheAspGluLysAsnIle 112
Db 247 AATATTGAGGCACCATCTCTGATACCCCTTAACGCGCAACAATATGATGAGAAATATA 306

Qy 113 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGlu 132
Db 307 CGACGAAGAGCTTATGATGCTTTGAATGTTCTGATGGCTATGACATTTATCTAAAGAT 366

Qy 133 LysIleGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeu 152
Db 367 AAAAAGAGATCCAGTGGAGGGCTTGCCCGGT-----ACTAGTATAAGTGACATT 417

Qy 153 GluGluValArgLysGluLeuVal-----AsnLysIleArgAsnLysLysAlaLeu 169
Db 418 GAAGAATGAGACGGAGCTTGTGGACTGAAGCTAGGATTGAAAGAAAGTGCTTAT 477

Qy 170 LeuGlnGluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsn---Gln 188
Db 478 CTACAGGAGCTACAAGACCAATATGTAGGTCTGCAGAAACCTGATTCACCGAATGAACAG 537

Qy 189 ThrLeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLys 208
Db 538 TCATATGGTTTCAGAAACACGCGCTTCTGGTGGAGTGGCTTTGCCATTTATCTTAGTTCAG 597

Qy 209 ThrSerArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHisPhe 228
Db 598 ACCGAGCTCATGCTACCGTGGAGATTGAGATATCAGAGATATGACAGCTGGTGCAATTT 657

Qy 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 658 GACTTCAATAGCACCCCATTCGAGCTGCACGACCTCATACGCTCCTTAAAGAAATGCGGA 717

Qy 249 -----ArgAsnSerIleGlyArgAlaGlyArgAla 258
Db 718 TTCTGTGAAGAGAACCAATGACAGCACTCAAGAGTCGATATCAAAATGGAGGTGAGAGC 777

Qy 259 Thr 259
Db 778 TCA 780
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 6, 2005, 21:42:57 ; Search time 196 Seconds  
(without alignments)

2178.920 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPPRGAAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFTO.spool/US1008830/runat\_04022005.122603.29577/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1008830 @CGN\_1.1.59 @runat\_04022005.122603.29577 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	399.5	30.3	2320	4	US-09-023-655-1456
2	389.5	29.5	1157	2	US-08-723-415B-5
3	389.5	29.5	1157	3	US-09-189-627A-5
4	389.5	29.5	1157	3	US-09-710-861-5
5	389	29.5	1202	2	US-08-723-415B-7
6	389	29.5	1202	3	US-09-189-627A-7
7	389	29.5	1202	3	US-09-710-861-7
8	389	29.5	1385	2	US-08-723-415B-1
9	389	29.5	1385	3	US-09-189-627A-1
10	389	29.5	1385	3	US-09-710-861-1
11	385	29.2	1427	4	US-09-949-016-3349
12	384	29.1	1154	2	US-08-723-415B-3

13	384	29.1	1154	3	US-09-189-627A-3	Sequence 3, Appli
14	384	29.1	1154	3	US-09-710-861-3	Sequence 3, Appli
15	378.5	28.7	1700	2	US-08-428-131-1	Sequence 1, Appli
16	378.5	28.7	1700	2	US-08-602-846-1	Sequence 1, Appli
17	378.5	28.7	1700	3	US-09-078-596-1	Sequence 1, Appli
18	374.5	28.4	1440	4	US-09-023-655-1399	Sequence 1399, Ap
19	374.5	28.4	1672	4	US-09-949-016-2937	Sequence 2937, Ap
20	261.5	19.8	358	4	US-09-640-211A-566	Sequence 566, App
21	254.5	19.3	364	4	US-09-640-211A-465	Sequence 465, App
22	187.5	14.2	352	4	US-09-640-211A-440	Sequence 440, App
C 23	156	11.8	601	4	US-09-949-016-120497	Sequence 120497,
C 24	156	11.8	601	4	US-09-949-016-120498	Sequence 120498,
25	156	11.8	193169	4	US-09-949-016-15091	Sequence 15091, A
26	155.5	11.8	58273	4	US-09-949-016-14679	Sequence 14679, A
27	155	11.8	601	4	US-09-949-016-105247	Sequence 105247,
28	146.5	11.1	271	4	US-09-640-211A-576	Sequence 576, App
29	128.5	9.7	1314	3	US-09-242-737-1	Sequence 1, Appli
30	128.5	9.7	2456	1	US-07-882-711-1	Sequence 1, Appli
31	128.5	9.7	2456	2	US-08-462-174-1	Sequence 1, Appli
32	128.5	9.7	2457	2	US-08-723-415B-12	Sequence 12, Appl
33	128.5	9.7	2457	3	US-09-189-627A-12	Sequence 12, Appl
34	128.5	9.7	2457	3	US-09-710-861-12	Sequence 12, Appl
35	128.5	9.7	2486	4	US-09-949-016-4380	Sequence 4380, Ap
36	128.5	9.7	2492	1	US-08-139-937-13	Sequence 13, Appl
37	128.5	9.7	2492	4	US-09-023-655-1230	Sequence 1230, Ap
38	128.5	9.7	2492	5	PCT-US93-11310-13	Sequence 13, Appl
39	128.5	9.7	2517	1	US-08-306-691B-18	Sequence 18, Appl
40	128.5	9.7	2517	1	US-08-385-142-2	Sequence 2, Appli
41	128.5	9.7	2517	2	US-08-481-814A-1	Sequence 1, Appli
42	128.5	9.7	2517	3	US-08-801-092-2	Sequence 2, Appli
43	128.5	9.7	2517	3	US-09-517-584A-3	Sequence 3, Appli
44	128.5	9.7	2517	3	US-09-315-113-2	Sequence 2, Appli
45	123.5	9.4	2078	4	US-09-799-451-486	Sequence 486, App

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-1456  
; Sequence 1456, Application US/09023655  
; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESS: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1456:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9604478
US-09-023-655-1456

Alignment Scores:
Pred. No.: 6.43e-39 Length: 2320
Score: 399.50 Matches: 94
Percent Similarity: 58.33% Conservative: 53
Best Local Similarity: 37.30% Mismatches: 88
Query Match: 30.29% Indels: 17
DB: 4 Gaps: 7

US-10-088-830-2 (1-261) x US-09-023-655-1456 (1-2320)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
Db 222 GTTACTCAGACACATAGCAGAGCTACTGGCTGGTCCCT-----GGT 266
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 267 GATAGAAAACGGCTAGAAAATTTATAGACTCTGATTTTTCAGAAAGTAAACGACAA 326
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 327 AAAGGAGATAAA---AATGGGAAGGCTTGAGACACTTTTTCATGAAAGTGTGTGAGAA 383
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 384 GTTCAAGAAAGTCAACATCGTACATGAATGCTGATGAGCTGGTGCAGAGTTC 443
QY 98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
Db 444 ACCAATTCAAATAACCAATTTGGCTGCTGATTCGAGGCTTATGATCAGAAAGACATTAGG 503
QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db 504 CGAAGAGTTTATGCTTTTAAATGTGCTAATGAATGCTGATGAGCTGGTGCAGAGTTC 563
QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db 564 AAAGAAATCAAGTGGATGGCTTGGCTGCCCTACCAATTCGCTCAGGAATGTCAGAACTGGAG 623
QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
Db 624 ATAGAGAAGCAGAGCGGATAGAACGATAAAGCAGAGCGGCGCCAGCTGCAAGAACTT 683
QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db 684 CTCTCAGCAAAATCGCTTTCAAAAACCTGGTGCAGAGAAATCGCAAAATGAGAGCAA 743
QY 194 AlaGluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThr 209
Db 744 AACAGGGCCCGCGGCTCTGAACCTCTACATTCAGCTGCCATTCATATCATCAATACA 803
QY 210 SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---Phe 228
Db 804 AGCAGAAAAACAGTCATAGATTCGAGCATCTCCAGTGAC---AAGTTTGATGATCTTTTC 860
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 861 AATTTTGACAAACACC---TTTGAGATTCATGATGACATAGATGACTAAAGCGGATGGGA 917
QY 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 917
```

```
Db 918 ATGTCGTTGGCTGGAGTCAGGCAAAATGCTCTCTG 953

RESULT 2
US-08-723-415B-5
; Sequence 5, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1110
; US-08-723-415B-5

Alignment Scores:
Pred. No.: 3.73e-38 Length: 1157
Score: 389.50 Matches: 94
Percent Similarity: 57.54% Conservative: 51
Best Local Similarity: 37.30% Mismatches: 74
Query Match: 29.53% Indels: 33
DB: 2 Gaps: 8

US-10-088-830-2 (1-261) x US-08-723-415B-5 (1-1157)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGACTGAGCTGCT----- 111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GCTGGGTTCCAGTAAACGAAAGCAAGAGATAAA--- 150
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGGAAGGCTTGAGACATTTTTCATGAAGGTGTGTGAGAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
```

196 GTTCAGCGGAAAGGCACAACTTCATCAATAGAGTACCTGATGAGTGGTATCTGAGTTT 255  
Db  
98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113  
Qy  
256 ACCAACTCAAAATACCATCTGGCAGCTGATTCGAGGCTTATGATCAGGAGAACATTAGA 315  
Db  
114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133  
Qy  
316 CGAAGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAGAGAA 375  
Db  
134 LysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153  
Qy  
376 AAAGAATAACAAGTGGATGGCTGCCTACCAATTCCTCAGGAATGCCAGAACCTGGAA 435  
Db  
154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173  
Qy  
436 ATCGAAGACGAGCGGATAGAACGGATAAACGAGCGAGCCGAGCTTACAGAACTT 495  
Db  
174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193  
Qy  
496 CTCCTTCAGCAAAATGCTTTTAAACCTGGTACAGCAATTCGAGGAAATCGACAAACAA 555  
Db  
194 AlaGluAsn-----ValAsnGly-----IleArgLeuProPheValLeuValArg 248  
Qy  
616 AGCAGAAACACGTCATGAGTCCAGCATCTCCAGTGAC---AAATTTGAATACCTTTT 672  
Db  
229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248  
Qy  
730 ATGTCCTTTGGTCTGGAGTCAGGCAAAATGCTCTCTG 765  
Db

RESULT 3

US-09-189-627A-5  
; Sequence 5, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1157  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1110)  
US-09-189-627A-5

Alignment Scores:  
Pred. No.: 3,73e-38 Length: 1157  
Score: 389,50 Matches: 94  
Percent Similarity: 57,54% Conservative: 51  
Best Local Similarity: 37,30% Mismatches: 74  
Query Match: 29,53% Indels: 33  
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x US-09-189-627A-5 (1-1157)  
Qy 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37  
Db 82 GTCACTCAGACTCACATAGCTGAGGCTGCT----- 111  
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57  
Db 112 -----GGCTGGGTTCAGTAAACGAAGCAAAAGAGAGATAA--- 150  
Qy 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77  
Db 151 -----AATGGGAAGGCTTCAGACATTTTTCATGAAGGTGTGTGAGAA 195  
Qy 78 ValGluAlaLysGlyArgThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97  
Db 196 GTTCAGCGGAAAGGCACAACTTCATCAATAGAGTACCTGATGAGTGGTATCTGAGTTT 255  
Qy 98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113  
Db 256 ACCAACTCAAAATACCATCTGGCAGCTGATTCGAGGCTTATGATCAGGAGAACATTAGA 315  
Qy 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133  
Db 316 CGAAGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAGAGAA 375  
Qy 134 LysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153  
Db 376 AAAGAATAACAAGTGGATGGCTGCCTACCAATTCCTCAGGAATGCCAGAACCTGGAA 435  
Qy 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173  
Db 436 ATCGAAGACGAGCGGATAGAACGGATAAACGAGCGAGCCGAGCTTACAGAACTT 495  
Qy 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193  
Db 496 CTCCTTCAGCAAAATGCTTTTAAACCTGGTACAGAAATCGACAAATGAACAA 555  
Qy 194 AlaGluAsn-----ValAsnGly-----IleArgLeuProPheValLeuValArg 209  
Db 556 AACGAGCGCTCCAGCTGTGAAATTCACCAATTCAGTCCATTTATAATCATTAATACA 615  
Qy 210 SerArgLysAlaArgValGluIleGluLysSerAspSerLysPheAlaHis---Phe 228  
Db 616 AGCAGAAACACGTCATGAGTCCAGCATCTCCAGTGAC---AAATTTGAATACCTTTT 672  
Qy 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248  
Db 673 AAATTTGATAACACC---TTTGAGATCCAGCAGCATAGAGGTACTGAAGCGGATGGGA 729  
Qy 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260  
Db 730 ATGTCCTTTGGTCTGGAGTCAGGCAAAATGCTCTCTG 765  
Db

RESULT 4

US-09-710-861-5  
; Sequence 5, Application US/09710861  
; Patent No. 6387649  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/710,861  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: US/09/189,627  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0



```
; SEQ ID NO 5
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1110)
US-09-710-861-5

Alignment Scores:
Pred. No.: 3,73e-38 Length: 1157
Score: 389.50 Matches: 94
Percent Similarity: 57.54% Conservative: 51
Best Local Similarity: 37.30% Mismatches: 74
Query Match: 29.53% Indels: 33
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x US-09-710-861-5 (1-1157)

Qy 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGATGAGGCTGCT----- 111
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GGCTGGGTCCAGTAACGAAGCAAAAGAGATATAA--- 150
Qy 58 AlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGAAAGGCTTGAGACATTTTCAATGAAGGTGTGTGAGAA 195
Qy 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 196 GTTCAGCGGAAGGCAACCTTCATCAATGAGTAGCTGATGAGCTGGTATCTGAGTTT 255
Qy 98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
Db 256 ACCAACTCAATAACCATCTGGCAGCTGATTCGACGGCTTATGATCAGGAGACATTAGA 315
Qy 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db 316 CGAAGAGTTTATGATGCTTTAAATGTTACTAATGCGCATGAACATAATTTCAAAGGAAAA 375
Qy 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db 376 AAGAATAACAGTGGATTGGCTTCCCTACCAATTCCTCAGGAATGCCAGAACTTGGAA 435
Qy 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
Db 436 ATCGAGAAGCAGAGCGGATAGAACGATAGAACGAGAGCGGCCAGCTTACAGAACTT 495
Qy 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db 496 CTCCTTCAGCAAAATGCTTTTAAAAACCTGGTACAGAGAATCGCAAAATGAACAACA 555
Qy 194 AlaGluAsn-----ValAsnGly-----IleArgLeuProPheValLeuValTyr 209
Db 556 AACCAGGCGCCCTCAGCTGGAATTCACCATTCAGCTGCATTTATATCATTAATACA 615
Qy 210 SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---Phe 228
Db 616 AGCAGGAAACAGTCATAGACTCAGCATCTCCAGTGAC---AAATTTGATACCTTTT 672
Qy 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluValArg 248
Db 673 AATTTTGAATACACC---TTTGATCCACGACATAGAGGTACTGAGCGGATGGA 729
Qy 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 730 ATGTCCTTGGTCTGGAGTCAGGCAATGCTCTCTG 765

US-10-088-830-2 (1-261) x US-08-723-415B-7 (1-1202)

Qy 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGATGAGGCT-----GCTGGCTGG 117
Qy 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
Db 118 GTTCCAGGTAGTAAAGAACGAGTAGAGATTTATAGACTCTGATTTTTCAGAAAGTAA 177
Qy 55 GluLysAlaAlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIleVal 74
Db 178 CGAAGCAAAAGGAGGATATAA---AATGGAAAGGCTTGAGACATTTTCAATGAAGGTG 234
Qy 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
Db 235 TGTGAAAAGTTTCAGCGGAAAGGCAACCTTCATCAATGAGTAGCTGATGAGCTGGTA 294
Qy 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
```





FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1155)  
US-09-710-861-7

## Alignment Scores:

Pred. No.: 4,57e-38 Length: 1202  
Score: 389.00 Matches: 98  
Percent Similarity: 57.48% Conservative: 48  
Best Local Similarity: 38.58% Mismatches: 86  
Query Match: 29.49% Indels: 22  
DB: 3 Gaps: 10

US-10-088-830-2 (1-261) x US-09-710-861-7 (1-1202)

```

Qy 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGATAGCTGAGCT-----GCTGGCTGG 117
Qy 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
Db 118 GTTCCCAAGTAGTAAACAGCTAGAGATTATAGACTCTGATTTTTCAGAAAGTAA 177
Qy 55 GluLysAlaAlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleVal 74
Db 178 CGAAGCAAAAAGAGAGATAA---AATGGGAAGGCTTGAGACATTTTCAATGAAGTG 234
Qy 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
Db 235 TGTGAGAAAGTTTCAAGCGAAAGGCACAACTTCATACATGAGGTAGCTGATGAGCTGTA 294
Qy 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 295 TCTGAGTTTACCAACTCAATACCATCTGCGAGCTGATTCGGCTTATGATCAGGAGAC 354
Qy 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
Db 355 ATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAG 414
Qy 132 GluLysLysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 415 GAAAAAAGAAATCAAGTGGATGGCTGCTCCCTACCAATTTCTGCTCAGGAATGCCAGAAC 474
Qy 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGln 171
Db 475 CTGGAATCAGAGAGCAGCGGATAGACGGAATAGCAGAAAGCAGAGCCAGCTACAA 534
Qy 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLysLeuArgAsnGlnThrLeuGlu 191
Db 535 GAATCTCTCTTCAGCAAAATGCTTTTAAACCTGTGTACAGAGAAATCGACAAATGAA 594
Qy 192 SerSerAlaGluAsn-----ValAsnGly---IleArgLeuProPheValLeuVal 207
Db 595 CAACAAACACGAGGCCCTCCAGCTGTGAATTCACCAATTCAGCTGCCATTTATAATCAT 654
Qy 208 LysThrSerArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHis 227
Db 655 AATACAGAGGAAACAGATCAGTACTGCGCATCTCCAGTAC---AATTTGAATAC 711
Qy 228 ---PheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGly 246
Db 712 CTTTTTAATTTGATAACACC---TTTGAGATCCAGCCGACATAGAGGTACTGAAGCG 768
Qy 247 ValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 769 ATGGGAATGCTCTTTGCTGGAGTCAGGCAATGCTCTCTG 810

```

## RESULT 8

US-08-723-415B-1  
Sequence 1, Application US/08723415B  
Patent No. 5853199  
GENERAL INFORMATION:  
APPLICANT: Lathangue, Nicholas B.

```

APPLICANT: delaluna, Susana
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 No. 5859199th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,415B
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610195.1
FILING DATE: 15-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 117-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1338
US-08-723-415B-1

Alignment Scores:
Pred. No.: 5,7e-38 Length: 1385
Score: 389.00 Matches: 98
Percent Similarity: 57.48% Conservative: 48
Best Local Similarity: 38.58% Mismatches: 86
Query Match: 29.49% Indels: 22
DB: 2 Gaps: 10

US-10-088-830-2 (1-261) x US-08-723-415B-1 (1-1385)
Qy 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
Db 265 GTCACTCAGACTCAGATAGCTGAGCT-----GCTGGCTGG 300
Qy 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
Db 301 GTTCCCAAGTAGTAAACAGCTAGAGATTATAGACTCTGATTTTTCAGAAAGTAA 360
Qy 55 GluLysAlaAlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleVal 74
Db 361 CGAAGCAAAAAGAGAGATAA---AATGGGAAGGCTTGAGACATTTTCAATGAAGTG 417
Qy 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
Db 418 TGTGAGAAAGTTTCAAGCGAAAGGCACAACTTCATACATGAGGTAGCTGATGAGCTGTA 477
Qy 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 478 TCTGAGTTTACCAACTCAATACCAATACCTGCGAGCTGATTCGGCTTATGATCAGGAGAC 537
Qy 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131

```

Db 538 ATTAGACGAGAGCTTTATGCTTTAAATGTACTAATGGCGATGAACATAATTTCAAG 597  
Qy 132 GluLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 151  
Db 598 GAAAAAAGAAATCAAGTGGATGGCTGCTTACCAATTTCTCTCAGGAATGCCAGAAC 657  
Qy 152 LeuGluGluValArgLysGluLeuValAsnLysLeuLeuLeuLeuLeuLeuLeuLeuGln 171  
Db 658 CTGGAATCGAGAGCAGAGCGGATAGAACCGGATAAGCAGAGCGAGCCAGCTACAA 717  
Qy 172 GluLeuGluLysGlnPheAspLeuGlnAsnLysLeuLeuLeuLeuLeuLeuLeuLeu 191  
Db 718 GAACCTCTCTCAGCAAAATGCTTTTAAACCTGCTACAGAGAAATCGCAAAATGAA 777  
Qy 192 SerSerAlaGluAsn-----ValAsnGly---IleArgLeuProPheValLeuVal 207  
Db 778 CAACAAACACCGGCGCTCAGCTGGAATTCACCAATTCAGCTGCAATTTATAATCATT 837  
Qy 208 LysThrSerArgLysAlaArgValGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 227  
Db 838 AATACAGCAGAGAAACAGCTAGACTGAGCTGAGCTCTCCAGTCAC---AAATTTGAATAC 894  
Qy 228 ---PheGluPheAsnGlyAlaProPheThrLeuLeuLeuLeuLeuLeuLeuLeuGly 246  
Db 895 CTTTAAATTTGATAACACC---TTTGAGATCCAGCAGACATAGAGGTACTGAAAGCG 951  
Qy 247 ValArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260  
Db 952 ATGGGAATGCTCTTGGTCTGGAGTCAGGCAATGCTCTG 993

## RESULT 9

US-09-189-627A-1  
; Sequence 1, Application US/09189627A  
; Patent No. 6159691  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/189, 627A  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1385  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1338)  
US-09-189-627A-1

## Alignment Scores:

Pred. No.:	5,7e-38	Length:	1385
Score:	389.00	Matches:	98
Percent Similarity:	57.48%	Conservative:	48
Best Local Similarity:	38.58%	Mismatches:	86
Query Match:	29.49%	Indels:	22
DB:	3	Gaps:	10

US-10-088-830-2 (1-261) x US-09-189-627A-1 (1-1385)

Qy 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37  
Db 265 GTCACCTCAGACTCAGCTAGCTGAGCT-----GCTGCTGG 300  
Qy 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54

Db 301 GTTCCAGTAGTAGAAACGAGCTAGAGAAATTTATAGACTCTGATTTTTCAGAAAGTAAA 360  
Qy 55 GluLysAlaAlaProArgIleThrGlyTTPGlyLeuArgGluTyrSerLysIleVal 74  
Db 361 CGAAGCAAAAAAGGAGATAAA---AATGGAAAGGCTTGAGACATTTTCAATGAAGGTG 417  
Qy 75 CysGluLysValLeuLysGlyArgThrThrTyrAsnGluValAlaLeuLeuLeuLeuTyr 94  
Db 418 TGTGAGAAAGTTTCAGCGGAAAGCACAACCTTCATCAATCAGGTAGCTGATGAGCTGTA 477  
Qy 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111  
Db 478 TCTGAGTTTACCAACTCAATTAACCATCTGGCAGCTGATTCGGCTTTATGATCAGGAGAAC 537  
Qy 112 IleArgArgArgValTyrAspAlaPheAsnValLeuLeuLeuLeuLeuLeuLeuLeuLys 131  
Db 538 ATTAGACGAGAGTTTATGCTTTAAATGTACTAATGGGATGAACATAATTTCAAG 597  
Qy 132 GluLysLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 151  
Db 598 GAAAAAAGAAATCAAGTGGATGGCTGCTTACCAATTTCTCTCAGGAATGCCAGAAC 657  
Qy 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGln 171  
Db 658 CTGGAATCGAGAGCAGAGCGGATAGAACCGGATAAGCAGAGCGAGCCAGCTACAA 717  
Qy 172 GluLeuGluLysGlnPheAspLeuGlnAsnLysLeuLeuLeuLeuLeuLeuLeuLeu 191  
Db 718 GAACCTCTCTCAGCAAAATGCTTTTAAACCTGCTACAGAGAAATCGCAAAATGAA 777  
Qy 192 SerSerAlaGluAsn-----ValAsnGly---IleArgLeuProPheValLeuVal 207  
Db 778 CAACAAACACCGGCGCTCAGCTGGAATTCACCAATTCAGCTGCAATTTATAATCATT 837  
Qy 208 LysThrSerArgLysAlaArgValGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 227  
Db 838 AATACAGCAGAGAAACAGCTAGACTGAGCTGAGCTCTCCAGTCAC---AAATTTGAATAC 894  
Qy 228 ---PheGluPheAsnGlyAlaProPheThrLeuLeuLeuLeuLeuLeuLeuLeuGly 246  
Db 895 CTTTAAATTTTGAATAACACC---TTTGAGATCCAGCAGACATAGAGGTACTGAAAGCG 951  
Qy 247 ValArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260  
Db 952 ATGGGAATGCTCTTGGTCTGGAGTCAGGCAATGCTCTG 993

## RESULT 10

US-09-710-861-1  
; Sequence 1, Application US/09710861  
; Patent No. 6387649  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas  
; APPLICANT: de la Luna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
; FILE REFERENCE: 620-54  
; CURRENT APPLICATION NUMBER: US/09/710, 861  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: US/09/189, 627  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 08/723,415  
; PRIOR FILING DATE: 1996-09-30  
; PRIOR APPLICATION NUMBER: GB 9610195  
; PRIOR FILING DATE: 1996-05-15  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1385  
; TYPE: DNA  
; ORGANISM: mouse  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1338)  
US-09-710-861-1

```
Alignment Scores:
Pred. No.: 5,7e-38 Length: 1385
Score: 389.00 Matches: 98
Percent Similarity: 57.48% Conservative: 48
Best Local Similarity: 38.58% Mismatches: 86
Query Match: 29.49% Indels: 22
DB: 3 Gaps: 10

US-10-088-830-2 (1-261) x US-09-710-861-1 (1-1385)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
DB 265 GTCACTCAGACTCAGACTGAGGCT-----GCTGGCTGG 300
QY 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
DB 301 GTTCCAGTAGTAGAAAACGAGCTAGAGATTTATAGACTCTGATTTTTCAGAAAGTAA 360
QY 55 GluLysAlaAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleVal 74
DB 361 CGAAGCAAAAAAGAGATAAA---AATGGGAAAGGCTTGAGACATTTTTCATCAAGGTG 417
QY 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
DB 418 TGTGAGAAAGTTTCAGCGGAAGGCACAACTTCATCAATGAGTAGCTGATGAGCTGTA 477
QY 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
DB 478 TCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAGAAC 537
QY 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
DB 538 ATTAGACGAAGCTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAAG 597
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
DB 598 GAAAAAAGAAATCAAGTGGATGGCTGCTCCCTACCAATTCCTGCTCAGGAAATGCCAGAAC 657
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGluGln 171
DB 658 CTGGAATTCGAGAAGCAGAGCGGATAGAACGGATAAAGCAGAACGAGCCAGCTACAA 717
QY 172 GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
DB 718 GAATCTTCTCTCAGCAAAATGCTTTTAAAAACCTGCTACAGAGAAATCGACAAAATGAA 777
QY 192 SerSerAlaGluAsn-----ValAsnGly---IleArgLeuProPheValLeuVal 207
DB 778 CAACAAACACGAGCCCTCAGCTGTAATTCACCAATTCAGTGCATTTATATCAAT 837
QY 208 LysThrSerArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHis 227
DB 838 AATACAAGCAGGAAACAGTCATAGACTGCGAGCATCTCCAGTCAC---AAATTTGAATAC 894
QY 228 ---PheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGly 246
DB 895 CTTTTTAAATTTTATAACACC---TTTGAGATCCAGCAGACATAGAGTACTGAAGCGG 951
QY 247 ValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
DB 952 ATGGGAATGTCCTTTGGTGGAGTCAGGCAAAATGCTCTCTG 993

RESULT 11
US-09-949-016-3349
; Sequence 3349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3349
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-3349
```

```
Alignment Scores:
Pred. No.: 1,87e-37 Length: 1427
Score: 385.00 Matches: 92
Percent Similarity: 58.82% Conservative: 48
Best Local Similarity: 38.66% Mismatches: 84
Query Match: 29.19% Indels: 14
DB: 4 Gaps: 8
```

US-10-088-830-2 (1-261) x US-09-949-016-3349 (1-1427)

```
QY 34 GluAlaGlyGlyAsnAlaValGlnArgLysGlyAla-----ValAspProAspLys 50
DB 336 GAATCTCAGTCCCAACAAAAAGTGTAGAAAACGGGCTAGAAAATTTATAGACTCTGATTT 395
QY 51 AspArgLysLysLysLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyr 70
DB 396 TCAGAAAGTAAACGAAGCAAAAAGAGAGATAAA---AATGGGAAAGGCTTGAGACACTTT 452
QY 71 SerLysIleValCysGluLysValGluAlaLysGlyArgThrTyrAsnGluValAla 90
DB 453 TCAATGAAAGTGTGTGAGAAAGTTCAACGAAAGGTACAACATCGTACAATGAAGTCGCT 512
QY 91 AspGluIleTyrSerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPhe 107
DB 513 GATGAGCTGGTGTGAGAGTTTCCAAATTCAAATAACCATTTGGCTGCTGATTCGGCTTAT 572
QY 108 AspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArg 127
DB 573 GATCAGAAACACATTAGCGAAGAGTTTATGATCTTAAATGCTTAATGCGCAATGAAC 632
QY 128 ValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGlu 147
DB 633 ATAATTTCAAAGGAAAAAAGAAATCAAGTGGATGGCTGCTGCTACCAATTCGCTCAG 692
QY 148 LysIleLysLysLysLysGluValArgLysGluLeuValAsnLysIleArgAsnLysLys 167
DB 693 GAATGTGAGATCTGGAGATAGAGAACGAGCGGATAGAACGGGATAAAGCAGAGACGG 752
QY 168 AlaLeuLeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn 187
DB 753 GCCCAGCTGCAGAACTTCTCTACAGCAAAATCGCTTTCAAAAACCTGGTACAGAGAAAT 812
QY 188 GlnThrLeuGluSerSerAlaGluAsnVal-----AsnGlyIleArgLeuPro 203
DB 813 CGAATAAATGAGCAGCAAAAACCGAGCGGCTCTGAACCTCTACCATTCAGCTGCCA 872
QY 204 PheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSerAspAspSer 223
DB 873 TTCAATATCATCAATCAAGCAAAAAACAGTCATGATTTGCAAGCATCTCCAGTCAG 929
QY 224 LysPheAlaHis---PheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSer 242
DB 930 AGTTTGGAGTATCTTTTCAATTTTGACACACC---TTTGAGATCTCATGATGACATGAA 986
QY 243 IleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
DB 987 GTACTAAAGCGGATGGATGTGCTTTGGCTGCTGAGTCAGGCAAAATGCTCTCTG 1040
```

RESULT 12

US-08-723-415B-3  
 ; Sequence 3, Application US/08723415B  
 ; Patent No. 5859199  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LaThangue, Nicholas B.  
 ; APPLICANT: delaluna, Susana  
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
 ; TITLE OF INVENTION: THEREOF  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NIXON & VANDERHYE P.C.  
 ; STREET: 1100 No. 5859199th Giebe Rd. 8th floor  
 ; CITY: Arlington  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22201-4741  
 ; COMPUTER READABLE FORM: disk  
 ; MEDIUM TYPE: Floppy  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/723,415B  
 ; FILING DATE: 30-SEP-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: GB 9610195.1  
 ; FILING DATE: 15-MAY-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Crawford, Arthur R.  
 ; REGISTRATION NUMBER: 25,327  
 ; REFERENCE/DOCKET NUMBER: 117-220  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-816-4000  
 ; TELEFAX: 703-816-4100  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1154 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1107  
 ; US-08-723-415B-3

Alignment Scores:  
 Pred. No.: 1.78e-37 Length: 1154  
 Score: 384.00 Matches: 94  
 Percent Similarity: 57.37% Conservative: 50  
 Best Local Similarity: 37.45% Mismatches: 75  
 Query Match: 29.11% Indels: 32  
 DB: 2 Gaps: 9

US-10-088-830-2 (1-261) x US-08-723-415B-3 (1-1154)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37  
 Db 82 GTCACTCAGACTCACATAGCTGAGGCTGCT-----111  
 QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGlyLysAla 57  
 Db 112 -----GCTCGGGTCCAGTAAACGAGCAAAAGAGATAAA---150  
 QY 58 AlaAlaProArgIleThrGlyLeuArgGlyLysSerLysIleValCysGluLys 77  
 Db 151 -----AATGGGAAGGCTTGAGACATTTTCATGAGGTGTGTGAGAA 195  
 QY 78 ValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThr 97  
 Db 196 GTTCAGCGGAAGGCACACTTCATACATATGAGGTAGCTGATGAGCTGATCTGAGTTT 255

QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114  
 Db 256 ACCAACTCAATAACCACTGGCAGCTGATTGGCTTATGATCAGAGAACATTAGACGA 315  
 QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134  
 Db 316 AGAGTTTATGATGCTTTAAATGCTACTAATGGCGATGAACATAATTTCAAAGGAAAAAAA 375  
 QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLysGlu 154  
 Db 376 GAAATCAAGTGGATTGGCTGCTTACCAATTCCTCAGAAATGCCAGAACCTGGAAATC 435  
 QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGlu 174  
 Db 436 GAGAAGCAGAGCGGATAGAACGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495  
 QY 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194  
 Db 496 CTTCAAGCAATTCCTTTTAAAAACCTGGTACAGAGAAATCGACAAATGAACAAACAA 555  
 QY 195 GluAsn-----ValAsnGly---IleArgLeuProPheValLeuValLysThrSer 210  
 Db 556 CAGGCGCTCCAGCTGTGAATTCACCATTCAGCTGCCATTTATAATCAATATACAAGC 615  
 QY 211 ArgLysAlaArgValGluIleSerAspSerLysPheAlaHis---PheGlu 229  
 Db 616 AGAAAAACAGTCATAGACTGCAGCATCTCCAGTGAC--AAATTTGAATACCTTTTAA 672  
 QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 249  
 Db 673 TTTGATAACACC--TTTGAGATCCAGCAGCAGCATAGAGGTACTGAAGCGGATGGA 729  
 QY 250 AsnSerIleGlyArgAlaGlyArgAlaThrLeu 260  
 Db 730 TCCTTTGGTCTCGAGTCAGGCAATGCTCTCTG 762

RESULT 13

US-09-189-627A-3  
 ; Sequence 3, Application US/09189627A  
 ; Patent No. 6159691  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Thangue, Nicholas  
 ; APPLICANT: de la Luna, Susana  
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
 ; FILE REFERENCE: 620-54  
 ; CURRENT APPLICATION NUMBER: US/09/189,627A  
 ; CURRENT FILING DATE: 1998-11-10  
 ; PRIOR APPLICATION NUMBER: 08/723,415  
 ; PRIOR FILING DATE: 1996-09-30  
 ; PRIOR APPLICATION NUMBER: GB 9610195  
 ; PRIOR FILING DATE: 1996-05-15  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1154  
 ; TYPE: DNA  
 ; ORGANISM: mouse  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(1107)  
 ; US-09-189-627A-3

Alignment Scores:  
 Pred. No.: 1.78e-37 Length: 1154  
 Score: 384.00 Matches: 94  
 Percent Similarity: 57.37% Conservative: 50  
 Best Local Similarity: 37.45% Mismatches: 75  
 Query Match: 29.11% Indels: 32  
 DB: 3 Gaps: 9

US-10-088-830-2 (1-261) x US-09-189-627A-3 (1-1154)

```
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 GTCACTCAGACTCAGACTCAGACTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
112 -----GGCTGGGTTCCCGATTAACGAGCAAAAGGAGATATA---150
QY 58 AlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 -----AATGGGAAAGGCTTGAGACATTTTCAATGAAGGTTGTGAGAAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 GTTCAGCGGAAGGCAACACTTCATCAATGAGTACGCTGAGCTGGTATCTGAGTTT 255
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 ACCAACTCAAATAACCATCTGGCAGCTGATTTCGGCTTATGATCAGGAGAACATTAGACGA 315
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 AGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAAGGAAAAA 375
QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 GAAATCAAGTGGATGGCTTCCCTACCAATTCGTCTCAGAAATGCCAGAACCTTGAATC 435
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GAGAAACAGTCATAGACTCGAGCATCTCCAGTGAC---AAATTTGAATACCTTTTAAAT 672
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArgArg 249
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
673 TTTGATAACACC---TTTGAGATCCACGACGACATAGAGTACTGAAGCGGATGGGAATG 729
QY 250 AsnSerIleGlyArgAlaGlyAlaThrLeu 260
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
730 TCCTTTGGTCTGGAGTCAGGCAAAATGCTCTCTG 762
```

## RESULT 14

```
US-09-710-861-3
; Sequence 3, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1154
```

```
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
US-09-710-861-3
```

## Alignment Scores:

```
Pred. No.: 1,78e-37 Length: 1154
Score: 384.00 Matches: 94
Percent Similarity: 57.37% Conservative: 50
Best Local Similarity: 37.45% Mismatches: 75
Query Match: 29.11% Indels: 32
DB: 3 Gaps: 9
```

US-10-088-830-2 (1-261) x US-09-710-861-3 (1-1154)

```
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 GTCACTCAGACTCAGACTCAGACTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
112 -----GGCTGGGTTCCCGATTAACGAGCAAAAGGAGATATA---150
QY 58 AlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 -----AATGGGAAAGGCTTGAGACATTTTCAATGAAGGTTGTGAGAAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 GTTCAGCGGAAGGCAACACTTCATCAATGAGTACGCTGAGCTGGTATCTGAGTTT 255
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 ACCAACTCAAATAACCATCTGGCAGCTGATTTCGGCTTATGATCAGGAGAACATTAGACGA 315
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 AGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAAGGAAAAA 375
QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 GAAATCAAGTGGATGGCTTCCCTACCAATTCGTCTCAGAAATGCCAGAACCTTGAATC 435
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GAGAAACAGTCATAGACTCGAGCATCTCCAGTGAC---AAATTTGAATACCTTTTAAAT 672
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArgArg 249
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
673 TTTGATAACACC---TTTGAGATCCACGACGACATAGAGTACTGAAGCGGATGGGAATG 729
QY 250 AsnSerIleGlyArgAlaGlyAlaThrLeu 260
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
730 TCCTTTGGTCTGGAGTCAGGCAAAATGCTCTCTG 762
```

## RESULT 15

```
US-08-428-131-1
; Sequence 1, Application US/08428131
; Patent No. 5863757
```

GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 5863757th Giebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 55..1284  
US-08-428-131-1

Alignment Scores:  
Pred. No.: 1.56e-36 Length: 1700  
Score: 378.50 Matches: 87  
Percent Similarity: 62.44% Conservative: 41  
Best Local Similarity: 42.44% Mismatches: 60  
Query Match: 28.70% Indels: 17  
DB: 2 Gaps: 7

US-10-088-830-2 (1-261) x US-08-428-131-1 (1-1700)

```
QY 50 LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 367 CGGAACAGGAGGGCGAGAG-----AATGGCAAGGGCCCTGCGGCAT 408

QY 70 TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 TTCTCCATGAAGGTGTGTGAGAGGTGCAGAGGAAGGACCACTCTCTACAAATGAGGTG 468

QY 90 AlaAspGluIleTyrSerGluLeuLysSerMet---AlaHisIle-----GlyGln 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 469 GCTGACGAGCTGTGGCAGAGTTTCAGCGCTGCAGACCAACCATCTTACCAACGAATCA 528

QY 106 GlyPheAspGluLysAsnIleArgArgArgValTyrAspAlaPheAsnValLeuIleAla 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 529 GCTTATGACGAGGAAGAACATCCGGCGCGGTGTCTACATGCCCTTAATGTCTAATGGCC 588

QY 126 LeuArgValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArg 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 589 ATGAACATCATCTCCAAGGAGAGAGGAGATCAATGGATCGCTGCCCACTCA 648

QY 146 TyrGluLysIleLysLysLeuGluValArgLysGluLeuValAsnLysIleArgAsn 165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 649 GCTCAGGAGTGCCAGAACTTAGAGGTGGAGGAGGAGGAGGAGGAGGATCAAAACAG 708
```

```
QY 166 LysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeu 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 709 AAGCAGTCTCAGCTCCAGGAGCTCATCTCGAGCAAAATTCCTTCAAGAACTTGGTGCAG 768

QY 186 ArgAsnGlnThrLeuGluSerSerAlaGluAsnVal-----AsnGlyIleArg 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 769 AGAAATCGCAAGCTGAGCAGCAGGCGCGCCCTCTCCCAACTCTGTCTATCCAC 828

QY 202 LeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSerAsp 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 829 TTGCCCTTCATCATTTGTCAACACCAGCAGACAGATCATTTGACTGCGACATCTCCAAT 888

QY 222 AspSerLysPheAlaHis---PheGluPheAsnGlyAlaProPheThrLeuHisAspAsp 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 889 GAC---AAATTTGAGTATCTGTTTAACTTTTGACAACACG---TTTGAGATCCACGATGAC 942

QY 241 LeuSerIleLeuGlu 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 943 ATTGAGGTGCTCAAG 957
```

Search completed: February 7, 2005, 00:24:07  
Job time : 201 secs

*This Page Blank (uspio)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 6, 2005, 23:22:07 ; Search time 580 Seconds  
(without alignments)  
2590.084 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10088830/runat\_04022005.122604\_29641/app\_query.fasta\_1.455  
-DB=Published Applications\_NA -QFMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10088830 @CGN 1 1 480 @runat\_04022005.122604\_29641  
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEREQUY -NEG\_SCORES=0 -WAIT -DSFBLCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTU5\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
C 1	842	63.8	1406	18	US-10-437-963-33888
2	779.5	59.1	774	18	US-10-425-115-4115
3	492	37.3	1639	18	US-10-739-930-1190
4	464	35.2	1350	17	US-10-425-114-4049
5	464	35.2	1483	18	US-10-425-115-2033
6	463	35.1	1585	18	US-10-425-115-87351
7	430	32.6	1431	17	US-10-424-599-43806
8	399.5	30.3	2320	17	US-10-641-643-1456
9	394	29.9	1266	9	US-09-954-456-1176
10	394	29.9	2968	18	US-10-357-930-20983
11	394	29.9	2968	18	US-10-357-930-20983
12	387	29.3	2968	14	US-10-198-846-10136
13	379.5	28.8	476	18	US-10-767-701-17642
14	379.5	28.8	1751	17	US-10-424-599-43105
15	374.5	28.4	692	15	US-10-206-901B-21
16	374.5	28.4	1440	17	US-10-172-118-1357
17	374.5	28.4	1440	17	US-10-160-554-4
18	374.5	28.4	1440	17	US-10-342-887-1357
19	374.5	28.4	1440	17	US-10-641-643-1399
20	374.5	28.4	1440	18	US-10-717-597-73
21	374.5	28.4	2607	15	US-10-106-698-569
22	372.5	28.2	1213	18	US-10-437-963-63675
23	370	28.1	642	18	US-10-767-795-6106
24	361.5	27.4	2673	16	US-10-240-965-234
25	361	27.4	1167	17	US-10-425-114-20051
26	354	26.8	1832	9	US-09-220-091-8
27	342.5	26.0	911	17	US-10-425-114-36537
28	337.5	25.6	792	18	US-10-437-963-64593
29	299.5	22.7	1690	15	US-10-053-248-23
30	299.5	22.7	1890	18	US-10-345-837-23
31	261.5	19.8	358	18	US-10-856-499-566
32	254.5	19.3	364	18	US-10-856-499-465
C 33	252.5	19.1	505	16	US-10-029-386-8935
C 34	227.5	17.2	690	11	US-09-969-034-3453
C 35	227.5	17.1	738	17	US-10-424-599-91931
36	213	16.1	774	18	US-10-425-115-88311
C 37	188.5	14.3	425	9	US-09-815-343-1294
C 38	188.5	14.3	425	17	US-10-097-105-1294
39	187.5	14.2	352	18	US-10-856-499-440
40	185.5	14.1	432	9	US-09-815-343-12
41	185.5	14.1	432	17	US-10-097-105-12
C 42	180	13.6	411	9	US-09-998-598-1532
43	167.5	12.7	488	17	US-10-242-535A-27279
44	167.5	12.7	488	17	US-10-085-783A-27279
45	146.5	11.1	271	18	US-10-856-499-576

# ALIGNMENTS

RESULT 1

US-10-437-963-33888/c

; Sequence 33888, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION: Thomas J.

; APPLICANT: La Rosa, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 33888

```
; LENGTH: 1406
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37957C.1
US-10-437-963-33888

Alignment Scores:
Pred. No.: 9,24e-79 Length: 1406
Score: 842.00 Matches: 185
Percent Similarity: 59.13% Conservative: 32
Best Local Similarity: 50.41% Mismatches: 36
Query Match: 63.84% Indels: 114
DB: 18 Gaps: 5

US-10-088-830-2 (1-261) x US-10-437-963-33888 (1-1406)

Qy 1 MetAlaProProArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaLeuLeuLeuThrGly 20
Db 1406 ATGCGCGCTCCCTCGCGGATGCGCGCGGCTGCCTCCGCGCGCGCGCGCTGGCCAA 1347

Qy 21 ValHisLeuGluAlaSerSerValProLeuProGlu----- 34
Db 1346 CTGCTCATTCGCGAAGCGCGCGCTCCCTTCGCGTCCGGAACGGTACCGCCCATTCGG 1287

Qy 34 ----- 34
Db 1286 CCCTGTACTTCTGATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1227

Qy 35 -----AlaGlyGlyAsnAlaValGlnArgLys----- 43
Db 1226 AAAAAATCGGTTCTTTACGCGCGCGCGCGCGCAATCGCGAGAGCGGAGGCG 1167

Qy 44 GlyAlaValAspProAspLysAspArgLysGluLysAlaAlaAlaProArgGlyThr 63
Db 1166 GGTGCG-----AACAGAACGGGAAGAGAACCGCGCGCGCAGAGATCAC 1116

Qy 64 GlyTrpGlyLeuArgGluTyrSerLysIle----- 73
Db 1115 GGGTGGGGGTCCTGAGTTCAGCAGATAGGGTTTACCAAACTGCCAGGCGCGGGTTA 1056

Qy 73 ----- 73
Db 1055 TCGCGCGCGCGGTAAAGCACAGTTACTGCGCGAGGTTACCGCGGTATTACATTACCG 996

Qy 73 ----- 73
Db 995 TACCGCGCGGTAAAGCGCGTAACCGCGGTAAACCGCGCGGTATCGGTGAAC 936

Qy 74 -----ValCysGluLysValGluAlaLysGlyArgThrThrTyrAsn 87
Db 935 CCTGGTGTGATTGCTTTGTTCTTAAAGAAAGTTAGGCCAAAGAGAACCATATATAT 876

Qy 88 Glu-----Val 89
Db 875 GAGATTATGTTCAACATCCAAAGATGAAGTCTATACAGCTCAGGAGAGCTTATAGTT 816

Qy 90 AlaAspGluLeuTyrSerGluLeuLysSerMetAlaHisIleGlyGlnGlyPheAspGlu 109
Db 815 GCGCATAGAGATTTTTCGGAGCTGAAGTCCATTACCGCAGAACCGTCTGGAGTTTGATG 756

Qy 110 LysAsnIleArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIle 129
Db 755 AAGATATTAGCGGAGGTATATGATGCTTCAATGTGCTCATTCGATTCGTGTTATT 696

Qy 130 AlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIle 149
Db 695 GCAAAAGATAAAAAGGAGATAAAGTGGATGGCTTACTAATATATAGATACGAAAAAGATA 636

Qy 150 LysLysLeuGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeu 169
Db 635 CAGAAAGTTGGAAGATTCAAGAAGAACTCATCACCAGGATCAAGAAGAAAGAGCTT 576
```

```
Qy 170 LeuGlnGluIleGlyLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThr 189
Db 575 CTCAGAAATTGAAAGCAGTTTGTGACCTTCAGATATTACATTAACCAACAGGCT 516

Qy 190 LeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLysThr 209
Db 515 AGTCAGAGCCGACAGAAAGTGTATGTCATCTCTCTCCGTTCTTATTGATCAAGACA 456

Qy 210 SerArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHisPheGlu 229
Db 455 TCCGAAAAAGCAAGGTTGAAATTCAGATTTCCGAAGATTCGAAGTTTCACGGTTTCG 396

Qy 230 PheAsnGlyAlaProPheThrIleHisAspAspLeuSerIleLeuGlyValArgArg 249
Db 395 TTCACGGTGCACCATTCACCATGTCATGATGATGATCAATCTTGAAGCCATCAGGCGT 336

Qy 250 AsnSerIleGlyArgAlaGly 256
Db 335 AACACAAAGAGAGAGCTGGC 315

RESULT 2
US-10-425-115-4115
; Sequence 4115, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 4115
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103754C.1
US-10-425-115-4115
```

```
Alignment Scores:
Pred. No.: 1,58e-72 Length: 774
Score: 779.50 Matches: 166
Percent Similarity: 77.73% Conservative: 33
Best Local Similarity: 64.84% Mismatches: 47
Query Match: 59.10% Indels: 10
DB: 18 Gaps: 4
```

US-10-088-830-2 (1-261) x US-10-425-115-4115 (1-774)

```
Qy 4 ProArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGlyValHisIle 23
Db 6 CCTCGCGCGATGCTGCGCGCGCAGCC-----GAGCTCACCAGCTCGGCATC 53

Qy 24 LeuGluAlaSerSerValProProLeuPro-----GluAlaGlyGlyAsnAlaVal 40
Db 54 AGCGAAGAGGCTGGGATCCCTCCGCTTCGCGACGCGGAAGCGGTTCGCGCAATGCGAGCT 113

Qy 41 GlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAlaAlaPro 60
Db 114 GAGGGAAGAGCGGCGGTATGAAGAACAAAGAGCGGAGGAGGTAAGCTGTGTACGCGAG 173

Qy 61 ArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAla 80
Db 174 CGGATCGCGGGTGGGGGCTCCGCGAGTTTCAGCAAGATAGTTTCTAAAAAGGTTCGAGCC 233

Qy 81 LysGlyArgThrThrThrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMet 100
Db 234 AAAGGACGCGACTTACATGAGTGGTTCAGATGAATTTTGTGGAGAGCTGAAG---TTG 290
```







[illegible]

## RESULT A

```

US-10-641-643-1456
:
: Sequence 1456, Application US/10641643
: Publication NO. US20040077003A1
: GENERAL INFORMATION:
: APPLICANT: Cocks, Benjamin G.
: Susan G. Stuart
: Jeffrey J. Seilhamer
: TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
: GENE EXPRESSION
:

```

ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/641,643  
 FILING DATE: 14-Aug-2003  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: <Unknown>  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1456:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2320 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9604478  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1456 :

US-10-641-643-1456

Alignment Scores:	7.74e-32	Length:	2320
Pred. No.:	399.50	Matches:	94
Score:	58.33%	Conservative:	53
Percent Similarity:	37.30%	Mismatches:	88
Best Local Similarity:	30.29%	Indels:	17
Query Match:	17	Gaps:	7
DB:	17		

US-10-088-830-2 (1-261) x US-10-641-643-1456 (1-2320)

QY	18	LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly	37
Db	222	GTCTACTCAGACACACATAGCAGAACTACTGGCTGGTCCGCT	266
QY	38	AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla	57
Db	267	GATAGAAAACGGGTAGAAAATTTATAGACTCTCAATTTTTCAGAAAGTAAACGAGCAAA	326
QY	58	AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys	77
Db	327	AAAGCAGATAAA--AATGGGAAGGCTTGAGACACTTTTCATCAAGTGTCTGAGAAA	383
QY	78	ValGluAlaLysGlyArgThrThyTyrAsnGluValAlaAspGluLleTyrSerGluLeu	97
Db	384	GTTCAGAAAGGTACAAACATCGTACAATGAAGTCGCTGATGAGCTGGTGTCTCAGAGTTC	443
QY	98	LysSer-----MetAlaHisIleGlyClnIleGlyPheAspGluLysAsnIleArg	113
Db	444	ACCAATTCAAATACCAATTTGGCTGCTGATTCGACGGCTTATGATCAGAAACATTAGG	503
QY	114	ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys	133
Db	504	CGAAGAGTTTATGATGCTTTAAATGTCTAATGCCAATGAACATAATTTCAAGAGAAATA	563
QY	134	LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu	153
Db	564	AAAGAAATCAAGTGGATTTGGCTCGCTTCCACCAATTTCTGCTCAGGAATGTCAGAACTCTGGAG	623
QY	154	GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle	173
Db	624	ATAGAAAGCAGAGCGGATAGAACCGATTAAGCAGAAAGCGGCCCGACGCTGCAAGAACTT	683
QY	174	GluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer	193
Db	684	CTCTCTACAGCAAAATCGCTTTTCAAAAACCTGGTCACAGAAATCGACAAATGAGCAGCAA	743
QY	194	AlaGluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThr	209
Db	744	AACCAAGCGCCCGCGGTCTGAACTTACCAATTCAGCTGCCATTCATAATCAATCAATACA	803
QY	210	SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---	228
Db	804	AGCGAAAAACAGTCATAGATTCGAGCATCTCCAGTGAC---AAGTTTGAGTATCTTTTC	860
QY	229	GluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg	248
Db	861	AATTTGACACACC---TTTGAGATCCATGATGACATAGAACTACTAAAGCGGATGGGA	917
QY	249	ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu	260
Db	918	ATGCTGTTGGCTTGGAGTCAGGCAAAATGCTCTCTG	953

## RESULT 9

```

US-09-954-456-1176
; Sequence 1176, Application US/09954456
; Patent No. US2002115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Ident
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76

```

```

; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; CURRENT APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1176
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1176

Alignment Scores:
Pred. No.: 1,29e-31 Length: 1266
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservatives: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: Gaps: 8

US-10-088-830-2 (1-261) x US-09-954-456-1176 (1-1266)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
DB 114 GTTACTCAGACACACATAGCAGAACTACTGGTGGTCCCT-----GGT 158
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
DB 159 GATAGAAAACGGCTAGAAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAAGCAA 218
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
DB 219 AAAGGAGATAAA---AATGGGAAGGCTTGACACATTTTCATGAAAGTGTGTGAGAA 275
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
DB 276 GTTCAACGAAAAGGTACAAATCGTACAATGAATCGCTGATGAGCTGGTGCAGAGTTC 335
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
DB 336 ACCAATTCAAATAACCATTTGGCTGCTGATTCGGCTTATGATCAGAAAGAACATTTAGGCGA 395
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
DB 396 AGAGTTTATGATGCTTTAATATGTCTTAATGGCAATGAACATATTTCAAGGAAAGAAAAA 455
QY 135 GluIleArgTrpMetClyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluGlu 154
DB 456 GAAATCAAGTGGCTGCTGCTCACTCAATTTCTGCTCAGGAATGTCTGAGATCTGGAGATA 515
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
DB 516 GAGAACGACAGCGGGATGAAAGCGATTAAGAGAGAGCGGGCCAGCTGCAAGAACTTC 575
QY 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194

US-10-357-930-20983
; Sequence 20983, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20983
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-20983

Alignment Scores:
Pred. No.: 4.1e-31 Length: 2968
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservatives: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: Gaps: 8

US-10-088-830-2 (1-261) x US-10-357-930-20983 (1-2968)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
DB 256 GTTACTCAGACACACATAGCAGAACTACTGGTGGTCCCT-----GGT 300
```



```
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLysAla 57
Db 301 GATAGAAAACGGCGTAGAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAGCAAA 360
Qy 58 AlaAlaProArgLysGlyTrpGlyLeuArgGluTyrSerLysLysValCysGluLys 77
Db 361 AAGGAGATATAA---AATGGAAAGGCTTGAGACACTTTTCAATGAAAGTGTGTGAGAA 417
Qy 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysLysSerGluLeu 97
Db 418 GTTCAACGAAAGGTACAACATCGTACAATGAATCGCTGATGAGTGTGTGAGAGTTC 477
Qy 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 478 ACCAATTCAATAACCAATTTGGCTGCTGATTCGGCTTATGATCAGAAACATTAGCGGA 537
Qy 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 538 AGAGTTTATGATGCTTTAAATGTGCTAATGCAATGAACATAATTTCAAGGAAAGAAAAA 597
Qy 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLysGluGlu 154
Db 598 GAAATCAAGTGGATGCTGCTGCTTACCAATTTCTGCTCAGGAATGTCAGAAATCTGGA 657
Qy 155 ValArgLysGluLeuValAsnLysLysAlaArgAsnLysLysAlaLeuGlnGluLeu 174
Db 658 GAGAGCAGAGCGGATAGAACGGATAAGCAAGAGAGCGGCGCCAGCTGCAAGAACTTCTC 717
Qy 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlnSerSerAla 194
Db 718 CTACAGCAAAATCGCTTTCAAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
Qy 195 GluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSer 210
Db 778 CAGGGCCCGCGGCTCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTG 837
Qy 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGlu 229
Db 838 AGAAAAACAGTCATAGATGTCAGACATCTCCAGTGAC---AAGTTTGAGTATCTTTCAAT 894
Qy 230 PheAsnGlyAlaProPheThrLeuHisAspLeuSerLysLysLysLysGluGluVal 249
Db 895 TTTGACAACACC---TTTGAGATCCATGATGATGATGATGATGATGATGATGATGATG 951
Qy 250 AsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 952 TCGTTTGGCTGGAGTCAGCAAAATGCTCTCTG 984

RESULT 11
US-10-357-930-26827
; Sequence 26827, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
```

```
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26827
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-26827

Alignment Scores:
Pred. No.: 4,1e-31 Length: 2968
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservative: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: 18 Gaps: 8

US-10-088-830-2 (1-261) x US-10-357-930-26827 (1-2968)
Qy 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 256 GTTACTCAGACACACATAGCAGAACTCTGCTGGTCCCT-----GGT 300
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 301 GATAGAAAACGGCTAGAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAGCAAA 360
Qy 58 AlaAlaProArgLysGlyTrpGlyLeuArgGluTyrSerLysLysValCysGluLys 77
Db 361 AAGGAGATATAA---AATGGAAAGGCTTGAGACACTTTTCAATGAAAGTGTGTGAGAA 417
Qy 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysLysSerGluLeu 97
Db 418 GTTCAACGAAAGGTACAACATCGTACAATGAATGCTGCTGCTGCTGCTGCTGCTGCTG 477
Qy 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 478 ACCAATTCAATAACCAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
Qy 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 538 AGAGTTTATGATGCTTTAAATGTGCTAATGCAATGAACATAATTTCAAGGAAAGAAAAA 597
Qy 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLysGluGlu 154
Db 598 GAAATCAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
Qy 155 ValArgLysGluLeuValAsnLysLysAlaArgAsnLysLysAlaLeuGlnGluLeu 174
Db 658 GAGAGCAGAGCGGATAGAACGGATAAGCAAGAGAGCGGCGCCAGCTGCAAGAACTTCTC 717
Qy 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlnSerSerAla 194
Db 718 CTACAGCAAAATCGCTTTCAAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
Qy 195 GluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSer 210
Db 778 CAGGGCCCGCGGCTCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTGAACTCTG 837
Qy 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGlu 229
Db 838 AGAAAAACAGTCATAGATGTCAGACATCTCCAGTGAC---AAGTTTGAGTATCTTTCAAT 894
Qy 230 PheAsnGlyAlaProPheThrLeuHisAspLeuSerLysLysLysLysGluGluVal 249
Db 895 TTTGACAACACC---TTTGAGATCCATGATGATGATGATGATGATGATGATGATGATG 951
```



```
QY 250 AenSerIleGlyArgAlaGlyArgAlaThrLeu 260
   ::::::::::::::::::::
Db 952 TCGTTGGCTGGAGTCAGGCAAAATGCTCTCTG 984

RESULT 12
US-10-198-846-10136
; Sequence 10136, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10136
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10136

Alignment Scores:
Pred. No.: 2,248-30 Length: 2968
Score: 387.00 Matches: 95
Percent Similarity: 58.33% Conservative: 52
Best Local Similarity: 37.70% Mismatches: 88
Query Match: 29.34% Indels: 17
DB: 14 Gaps: 8

US-10-088-830-2 (1-261) x US-10-198-846-10136 (1-2968)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
   ::::::::::::::::::::
Db 256 GTTACTCAGACACATAGACAGAGCTACTGGCTGGTCCCT-----GGT 300

QY 38 AenAlaValGlnArgLysGlyAlaValAseProAspLysArgLysLysGlyLysAla 57
   ::::::::::::::::::::
Db 301 GATAGAAAACGGGCTAGAAAATTTATAGACTCTGATTTTTCAGAAAAGTAAACGAAGCAAA 360

QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTySerLysIleValCysGluLys 77
   ::::::::::::::::::::
Db 361 AAGGAGATATAA---AATGGGAAGGCTTGAGACACTTTTCAATGAAGAGTGTGAGAAA 417

QY 78 ValGluAlaLysGlyArgThrThrTyArgGluValAlaAseGluIleTySerGluLeu 97
   ::::::::::::::::::::
Db 418 GTTCAACGAAAGGTACACATCGTACATGAATGAATGCTGATGAGCTGGTGTGAGTTTC 477

QY 98 ---LysSerMetAlaHisIleGly-----GlnGly-PheAspGluLysAseIleArg 114
   ::::::::::::::::::::
Db 478 ACCAATTCAATAAACCATTTGGCTGCTGATTCGGGCTTATGATCAGAAACATATTAGGCG 537

QY 114 GAGValTyArgAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
   ::::::::::::::::::::
Db 538 AAGAGTTTATGATGCTTTAAATGTGCTAATGCAATGAACATTAATTTCAAGGAAAAAAA 597

QY 134 sGluIleArgTyrMetGlyLeuSerAsnTyArgTyArgGluLysIleLysLysLeuGlu 154
   ::::::::::::::::::::
Db 598 AGAATCAAGTGGATTGGCTGCTGCTACCAATTCCTCAGGATGCTCAGAAATCTGGAGAT 657

QY 154 uValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle 174
   ::::::::::::::::::::
```

```
Db 658 AGAGAAGCAGAGCGCGGATAGAACGATAAAGCAGAGCGGCGCCAGCTGCAAGAACTTCT 717
QY 174 uLysGlnPheAspAspLeuGluAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAl 194
   ::::::::::::::::::::
Db 718 CCTACAGCAAAATCGCTTTCAAAACCTGGTACAGAAATCGACAAAATGAGCAGCAAAA 777

QY 194 aGluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSe 210
   ::::::::::::::::::::
Db 778 CCAGGCGCGCGCGCTCTGAACCTCTACCATTCAGCTGCCATTCAATCATCAATCAAG 837

QY 210 rArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheG 229
   ::::::::::::::::::::
Db 838 CAGAAAACAGTCATAGATTGCAGCATCTCCAGTGAC---AAGTTTTCAGTATCTTTTCAA 894

QY 229 uPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArgAr 249
   ::::::::::::::::::::
Db 895 TTTTGACAAACG---TTTGAGATCCATGATGATAGTAGTACTAAAGCGGATGGGAAT 951

QY 249 gAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
   ::::::::::::::::::::
Db 952 GTCGTTTGGCTGGAGTCAGGCAAAATGCTCTCTG 985

RESULT 13
US-10-767-701-17642
; Sequence 17642, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 17642
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-039-P1-K1-D2
US-10-767-701-17642

Alignment Scores:
Pred. No.: 1,168-30 Length: 476
Score: 379.50 Matches: 79
Percent Similarity: 84.21% Conservative: 17
Best Local Similarity: 69.30% Mismatches: 16
Query Match: 28.77% Indels: 2
DB: 18 Gaps: 1

US-10-088-830-2 (1-261) x US-10-767-701-17642 (1-476)
QY 139 MetGlyLeuSer-AsnTyArgTyArgGluLysIleLysLysLeuGluGluValArgLysG 158
   ::::::::::::::::::::
Db 3 ATGGGCTTGTCTACACTTCCATATGAATACATA---AAGTTGGAGAAATCTGCCAAGA 59

QY 158 uLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAs 178
   ::::::::::::::::::::
Db 60 ACTCATGATTAGGTTTAAAGCAAGAAAAAACTTCTCCAGGAAATTTGAACGACAGTTTGA 119

QY 178 pAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAs 198
   ::::::::::::::::::::
Db 120 TCACCTCCAGAAATCAAGTTTCGCAACCAAGTTTACTACAGAGCGCCAGCTGAGAGTCGGA 179

QY 198 nGlyIleArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleG 218
   ::::::::::::::::::::
Db 180 TGGTATCTGCTTCCATCTTATTGGTCAAGCATCCAGAAAAGCAAGGGTGAATTTGA 239

QY 218 uLysSerAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHi 238
   ::::::::::::::::::::
Db 240 GATTTTCAGAGAACTCAAGTTTGGTTCGACTTCAACTGTTACACCACTTCCACCTTGCA 299
```

QY 238 SASpAspLeuSerIleLeuGluGlyValArgArgAsnSer 251  
Db 300 TGACGATGTCCTCAATCTTGAAGCGATCAGGTGTAATAAT 339

RESULT 14  
US-10-424-599-43105  
; Sequence 43105, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 43105  
; LENGTH: 1751  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_138923C.1  
US-10-424-599-43105

Alignment Scores:  
Pred. No.: 6,76e-30 Length: 1751  
Score: 379.50 Matches: 96  
Percent Similarity: 47.89% Conservative: 40  
Best Local Similarity: 33.80% Mismatches: 63  
Query Match: 28.77% Indels: 86  
DB: 17 Gaps: 3

US-10-088-830-2 (1-261) x US-10-424-599-43105 (1-1751)

QY 50 LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu 69  
Db 299 AAAAAGAGCGGACACACCGGCTGGTGACGATGAAGAGTGTCTCCGCTAT 358

QY 70 TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89  
Db 359 TTTAGTATGAAGTGTGTGAGAGGTAGAAAGCAGGAGGAGATCATACATGAGGTG 418

QY 90 AlaAspGluIleTyrSerGlu----- 96  
Db 419 GCAGCTGAGCTTGTGCGAAGTGAATAATAATGTTTCCACAAGTTCATTATTGTATGGA 478

QY 96 ----- 96  
Db 479 AATGTTTCTGAAAACAGAAATCCAAACTGAGAGAAAAGAAACTGGTGAGAAATAAC 538

QY 96 ----- 96  
Db 539 GTCTTTCTTTGATGTCACAAATTCACAAAATTCGTGGTTGGATTCTCAATTTAGTA 598

QY 96 ----- 96  
Db 599 CATATACCTAGTTATTTTAAATAAATTCATATGGAATGTGTTATTTAGGTAGCTG 658

QY 97 -----LeuLysSerMetAlaHis-----IleGlyGln 105  
Db 659 ATGAACCTGTGGCGAATTTGCTGATCCCAATCAATGTTTCACTCTCGATCAG-CAA 717

QY 106 GlyPheAspGluLysIleArgArgValTyrAspAlaPheAsnValLeuIleAla 125  
Db 718 CAATATGATGAGAAAAAATTCCTCGAAGGTGTATGATGCTCTGAATGTTCTCATGGCA 777

QY 126 LeuArgValIleAlaLysGluLysGluIleArgTyrMetGlyLeuSerAsnTyrArg 145  
Db 778 ATGATATATTATTTCTAAGGAAAAAGGAAATTCATATGGAAGGGTCTCCCTCGTACTAGT 837

QY 146 TyrGluLysIleLysLysLysLeuGluGluValArgLysGluLeuValAsnLysIleArgAsn 165  
Db 838 CTAAGCGATATTGAGAGCTAAAGTCAGAGCGTCTTTGGGCTCAGGAATAGAAATGAAAG 897

QY 166 LysLysAlaLeuLeuGluGluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeu 185  
Db 898 AAAGCAGCCTATTGCAAGAGCTGGAGGAGCAATACATAGGTCTTCAGAACTTATTCAA 957

QY 186 ArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsn---GlyIleArgLeuProphe 204  
Db 958 CGAAATGAGCAACTATATAGCTCAGGAATGCTCCCAATGGAGGTGTATCTTTGCCCTTT 1017

QY 205 ValLeuValLysThrSerArgLysAlaAa-gValGluIleGluLysSerAspSerLys 224  
Db 1018 ATTCTGTGTCAGACACGCGCCCATGCAACTGTTGAAAGTGGAAATATCAGAAATATGCAG 1077

QY 225 PheAlaHisPheGluPheAsnGlyAlaPropheThrLeuHisAspAspLeuSerIleLeu 244  
Db 1078 CTTGTGCATTTTGATTATATTGCGCTCCCTTTGAGATGCATGATGATAACTATGTCTC 1137

QY 245 GluGlyValArg 248  
Db 1138 AAGCAATGAAA 1149

RESULT 15  
US-10-206-901B-21  
; Sequence 21, Application US/10206901B  
; Publication No. US20030100540A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG, ZHONGHUA  
; APPLICANT: DUBOIS, RAYMOND  
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES  
; FILE REFERENCE: VBLT:012US  
; CURRENT APPLICATION NUMBER: US/10/206,901B  
; PRIOR FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 60/308,370  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 692  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-206-901B-21

Alignment Scores:  
Pred. No.: 6,46e-30 Length: 692  
Score: 374.50 Matches: 86  
Percent Similarity: 62.44% Conservative: 42  
Best Local Similarity: 41.95% Mismatches: 60  
Query Match: 28.39% Indels: 17  
DB: 15 Gaps: 7

US-10-088-830-2 (1-261) x US-10-206-901B-21 (1-692)

QY 50 LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu 69  
Db 55 CGCAACAGGAAAGGAGAGAG-----AATGGCAAGGGCTACGGCAT 96

QY 70 TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89  
Db 97 TTCTTCATCAAGTCTTGGAGAGGTGCGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 156

QY 90 AlaAspGluIleTyrSerGluLeuLysSerMet---AlaHisIle-----GlyGln 105  
Db 157 GCAGACGAGCTGTTGCGGAGTTCAGTGTCTGCCGACCAACCATCTTACCAACAGGTCA 216

QY 106 GlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAla 125  
Db 217 GCTTATGACCAAAAACATAGACGGCGGTACGATGCTTAAACGTGCTAATGGCC 276

QY 126 LeuArgValIleAlaLysGluLysGluIleArgTyrMetGlyLeuSerAsnTyrArg 145

[illegible]

Search completed: February 7, 2005, 01:44:24  
Job time : 584 secs

1115 Page Blank (uspto)